

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
25 July 2002 (25.07.2002)

PCT

(10) International Publication Number
WO 02/057476 A2

- (51) International Patent Classification⁷: **C12P 13/02**, 7/42, C12N 15/52, 9/02, 15/03
- (74) Agents: **HANLEY, Elizabeth, A.** et al.; Lahive & Cockfield, LLP, 28 State Street, Boston, MA 02109 (US).
- (21) International Application Number: PCT/US02/01887
- (22) International Filing Date: 19 January 2002 (19.01.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/263,053 19 January 2001 (19.01.2001) US
- (71) Applicant (*for all designated States except US*): **OMNI-GENE BIOPRODUCTS INC.** [US/US]; 763(D) Concord Avenue, Cambridge, MA 02138 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): **HERMANN, Theron** [US/US]; 18 Chilhowie Drive, Kinnelon, NJ 07405 (US). **PATTERSON, Thomas, A.** [US/US]; 89 Church Street, North Attleboro, MA 02760 (US). **PERO, Janice, G.** [US/US]; 20 Solomon Pierce Road, Lexington, MA 02420 (US). **YOCUM, Roger, R.** [US/US]; 4 Orchard Lane, Lexington, MA 02420 (US). **BALDENIUS, Kai-Uwe** [DE/DE]; Kneippstrasse 16, 67063 Ludwigshafen (DE). **BECK, Christine** [DE/DE]; Max-Joseph-Strasse 35, 68167 Mannheim (DE).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:**
- without international search report and to be republished upon receipt of that report
 - with sequence listing part of description published separately in electronic form and available upon request from the International Bureau
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: METHODS AND MICROORGANISMS FOR THE PRODUCTION OF 3-(2-HYDROXY-3-METHYL-BUTYRYLAMINO)-PROPIONIC ACID (HMBPA)

(57) Abstract: The present invention features methods of producing 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid ("HMBPA") and α -hydroxyisovalerate (" α -HIV") utilizing microorganisms having modified pantothenate biosynthetic enzyme activities. Recombinant microorganisms and conditions for culturing same are also featured. Also featured are compositions including HMBPA and compositions including α -HIV.

WO 02/057476 A2

METHODS AND MICROORGANISMS FOR THE PRODUCTION OF 3-(2-HYDROXY-3-METHYL-BUTYRYLAMINO)-PROPIONIC ACID (HMBPA)

Related Applications

5 The present invention claims the benefit of prior-filed provisional Patent Application Serial No. 60/263,053, filed January 19, 2001 (pending). The present invention is also related to U.S. Patent Application Serial No. 09/667,569, filed September 21, 2000 (pending), which is a continuation-in-part of U.S. Patent Application Serial No. 09/400,494, filed September 21, 1999 (abandoned). U.S. Patent
10 Application Serial No. 09/667,569 also claims the benefit of prior-filed provisional Patent Application Serial No. 60/210,072, filed June 7, 2000, provisional Patent Application Serial No. 60/221,836, filed July 28, 2000, and provisional Patent Application Serial No. 60/227,860, filed August 24, 2000. The entire content of each of the above-referenced applications is incorporated herein by this reference.

15

Background of the Invention

Conventional means of synthesizing chemical compounds is *via* synthesis from bulk chemicals, a process which is limited by factors such as substrate availability and/or cost, difficulty in resolving complex mixtures of products, complexities in
20 synthesizing large quantities of compounds in purified form, and difficulty in producing chiral compounds. Accordingly, researchers have recently looked to bacterial or microbial systems that express enzymes useful for various biosynthetic processes, for example, in the synthesis of pharmaceutical compounds, research reagents, nutraceuticals, vitamins, nutritional supplements, antibiotic compounds and the like. In
25 particular, bioconversion processes have been evaluated as a means of favoring production of preferred compounds and recently methods of direct microbial synthesis have been the focus of much research in the areas of pharmaceuticals and agriculture.

Summary of the Invention

30 The present invention relates to a processes for the direct microbial synthesis of [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid or 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid ("HMBPA"), referred to interchangeably herein as " β -alanine 2-(*R*)-hydroxyisovalerate", " β -alanine 2-hydroxyisovalerate", " β -alanyl- α -hydroxyisovalerate", N-(2-hydroxy-3-methyl-1-oxobutyl)- β -alanine ("HMOBA")

and/or "pantothenate". In particular, it has been discovered that in microorganisms engineered to overexpress certain enzymes conventionally associated with pantothenate and/or isoleucine-valine (*ilv*) biosynthesis, an alternative biosynthetic pathway is present that competes for key precursors of pantothenate biosynthesis, namely α -ketoisovalerate (5) (α -KIV) and β -alanine. α -KIV is converted to α -hydroxyisovalerate (α -HIV) catalyzed by various reductase enzymes and α -HIV is subsequently condensed with β -alanine to produce HMBPA.

In one embodiment, the invention features a process for the production of 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA) that includes culturing 10 a microorganism having increased keto reductase activity or increased pantothenate synthetase activity in the presence of excess α -ketoisovalerate and excess β -alanine, such that HMBPA is produced. In another embodiment, the invention features a process for the production of 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA) that includes culturing a microorganism having increased keto reductase 15 activity and increased pantothenate synthetase activity in the presence of excess α -ketoisovalerate and excess β -alanine, such that HMBPA is produced. In one embodiment, the microorganism has a modified *panE* gene, for example, a modified *panE1* gene and/or a modified *panE2* gene (e.g., the *panE* gene is overexpressed, deregulated or present in multiple copies). In another embodiment, the microorganism 20 has a modified *panC* gene (e.g., the *panC* gene is overexpressed, deregulated or present in multiple copies). In another embodiment, the microorganism further has increased acetohydroxyacid isomeroreductase activity. In another embodiment, the microorganism is cultured under conditions of increased acetohydroxyacid isomeroreductase activity in the presence of excess α -ketoisovalerate and excess β - 25 alanine, such that HMBPA is produced. In yet another embodiment, the microorganism comprises a modified *ilvC* gene (e.g., the *ilvC* gene is overexpressed, deregulated or present in multiple copies). In yet another embodiment, the microorganism further has reduced ketopantoate hydroxymethyltransferase activity (e.g., has a modified *panB* gene, for example a *panB* gene that has been deleted).

30 In another aspect, the invention features a process for the production of 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA) that includes culturing a microorganism having reduced ketopantoate hydroxymethyltransferase activity in the presence of excess α -ketoisovalerate and excess β -alanine, such that HMBPA is produced. In another aspect, the invention features a method for enhancing production

of 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA) relative to pantothenate that includes culturing a recombinant microorganism under conditions such that the HMBPA production is enhanced relative to pantothenate production. In another aspect, the invention features a process for the production of 2-hydroxyisovaleric acid (α-HIV) that includes culturing a microorganism which overexpresses PanE1 or PanE2 and which further has reduced PanC or PanD activity under conditions such that α-HIV is produced. In another aspect, the invention features a process for the production of 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA) that includes culturing a recombinant microorganism having decreased expression or activity of *serA* or *glyA* under conditions such that HMBPA is produced. In another aspect, the invention features a process for the production of 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA) that includes culturing a recombinant microorganism having decreased expression or activity of *serA* and *glyA* under conditions such that HMBPA is produced. Conditions for culturing the above described microorganisms include, for example, conditions of increased steady state glucose, conditions of decreased steady state dissolved oxygen, and/or cultured under conditions of decreased serine. Products produced according to the above described processes and/or methods are also featured. Also featured are recombinant microorganisms utilized in the above-described methods.

Compounds produced according to the methodologies of the present invention have a variety of uses. For example, HMBPA can be used to synthesize inhibitors of HMG CoA Reductase (II) (Gordon *et al. Bio. Med. Chem. Lett.* 1(3):161 (1991). Inhibitors of HMG CoA Reductase (II) have been studied for use as in the treatment of hypercholesterolaemia and coronary atherosclerosis progression. Inhibitors of HMG CoA Reductase also have been used to reduce risk of cardiovascular events in patients at risk. Moreover, the HMBPA precursor 2-hydroxyisovalerate (α-HIV) has been demonstrated to have nutraceutical properties, for example, in the prevention of aging of the skin. In particular, α-hydroxy acids, such as α-HIV (or 2-hydroxyvaline), can be used to synthesize α-hydroxy esters which have been found to induce increased skin thickness by increasing biosyntheses of glycosaminoglycans, proteoglycans, collagen, elastin, and other dermal components. The compounds can be used to treat skin disorders such as age spots, skin lines, wrinkles, photoaging and aging.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

Brief Description of the Drawings

Figure 1 is a schematic representation of the pantothenate and isoleucine-valine (*ilv*) biosynthetic pathways. Pantothenate biosynthetic enzymes are depicted in bold and their corresponding genes indicated in italics. Isoleucine-valine (*ilv*) biosynthetic enzymes are depicted in bold italics and their corresponding genes indicated in italics.

Figure 2 is a schematic representation of the biosynthetic pathway leading to [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid ("HMBPA") in *B. subtilis*.

Figure 3 is a schematic depiction of the structure of [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid ("HMBPA").

Figure 4 is a HPLC chromatogram of a sample of medium from a 14 L fermentation of PA824.

Figure 5 is a mass spectrum depicting the relative monoisotopic mass of [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid.

Figure 6 depicts an alignment of the C-terminal amino acids from known or suspected PanB proteins.

Figure 7 is a schematic representation of the construction of the plasmid pAN624.

Figure 8 is a schematic representation of the construction of the plasmid pAN620.

Figure 9 is a schematic representation of the construction of the plasmid pAN636.

Figure 10 is a schematic representation of the construction of the plasmid pAN637 which allows selection for single or multiple copies using chloramphenicol.

Figure 11 is a schematic representation of the construction of the plasmid pAN238, a plasmid for overexpressing *B. subtilis* *panE2* from the P_{26} promoter.

Detailed Description of the Invention

The present invention is based, at least in part, on the discovery of a novel biosynthetic pathway in bacteria, namely the [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid ("HMBPA") biosynthetic pathway. In particular, it has been discovered that bacteria are capable of generating HMBPA from α -ketoisovalerate (α -KIV), a key product of the isoleucine-valine (*ilv*) biosynthetic pathway and precursor of the

pantothenate biosynthetic pathway. Production of HMBPA in bacteria involves at least the pantothenate biosynthetic enzymes ketopantoate reductase (the *panE1* gene product) and/or acetohydroxyacid isomeroreductase (the *ilvC* gene product) and results from the condensation of 2-hydroxyisovaleric acid (α -HIV), formed by reduction of α -KIV, and β -alanine, the latter reaction being catalyzed by the pantothenate biosynthetic enzyme pantothenate synthetase (the *panC* gene product). Production of HMBPA is achieved by increasing ketopantoate reductase (*e.g.*, PanE1) and/or PanE2 and/or acetohydroxyacid isomeroreductase activities (*e.g.*, IlvC) in microorganisms, for example, by overexpressing or deregulating the genes encoding said enzymes. Optimal production of HMBPA is achieved by decreasing or deleting ketopantoate hydroxymethyltransferase activity (the *panB* gene product) in microorganisms, for example, by modifying or deleting the *panB* gene which encodes ketopantoate hydroxymethyltransferase (*e.g.*, PanB), optionally in addition to increasing ketopantoate reductase and/or PanE2 and/or acetohydroxyacid isomeroreductase activities in said microorganisms. The substrates α -KIV and β -alanine are required for HMBPA production, the latter provided, for example, by β -alanine feeding and/or increased aspartate- α -decarboxylate activity (the *panD* gene product). Increasing substrate concentration (*i.e.*, α -KIV and/or β -alanine) further enhances production of HMBPA. α -KIV production can be increased by overexpressing *ilvBNCD* genes and/or *alsS*. HMBPA production can further be increased by limiting serine availability or synthesis in appropriately engineered microorganisms.

In order that the present invention may be more readily understood, certain terms are first defined herein.

The term "pantothenate biosynthetic pathway" includes the biosynthetic pathway involving pantothenate biosynthetic enzymes (*e.g.*, polypeptides encoded by biosynthetic enzyme-encoding genes), compounds (*e.g.*, precursors, substrates, intermediates or products), cofactors and the like utilized in the formation or synthesis of pantothenate. The term "pantothenate biosynthetic pathway" includes the biosynthetic pathway leading to the synthesis of pantothenate in microorganisms (*e.g.*, *in vivo*) as well as the biosynthetic pathway leading to the synthesis of pantothenate *in vitro*.

The term “pantothenate biosynthetic enzyme” includes any enzyme utilized in the formation of a compound (*e.g.*, intermediate or product) of the pantothenate biosynthetic pathway. For example, synthesis of pantoate from α -ketoisovalerate (α -KIV) proceeds *via* the intermediate, ketopantoate. Formation of ketopantoate is catalyzed by the pantothenate biosynthetic enzyme ketopantoate hydroxymethyltransferase (the *panB* gene product). Formation of pantoate is catalyzed by the pantothenate biosynthetic enzyme ketopantoate reductase (the *panE* gene product). Synthesis of β -alanine from aspartate is catalyzed by the pantothenate biosynthetic enzyme aspartate- α -decarboxylase (the *panD* gene product). Formation of pantothenate from pantoate and β -alanine (*e.g.*, condensation) is catalyzed by the pantothenate biosynthetic enzyme pantothenate synthetase (the *panC* gene product). Based on the newly discovered HMBPA biosynthesis pathway, pantothenate biosynthetic enzymes may also perform an alternative function as enzymes in the HMBPA biosynthetic pathway described herein.

The term “pantothenate” includes the free acid form of pantothenate, also referred to as “pantothenic acid” as well as any salt thereof (*e.g.*, derived by replacing the acidic hydrogen of pantothenate or pantothenic acid with a cation, for example, calcium, sodium, potassium, ammonium), also referred to as a “pantothenate salt”. The term “pantothenate” also includes alcohol derivatives of pantothenate. Preferred pantothenate salts are calcium pantothenate or sodium pantothenate. A preferred alcohol derivative is pantothenol. Pantothenate salts and/or alcohols of the present invention include salts and/or alcohols prepared *via* conventional methods from the free acids described herein. In another embodiment, calcium pantothenate is synthesized directly by a microorganism of the present invention. A pantothenate salt of the present invention can likewise be converted to a free acid form of pantothenate or pantothenic acid by conventional methodology.

The term “isoleucine-valine biosynthetic pathway” includes the biosynthetic pathway involving isoleucine-valine biosynthetic enzymes (*e.g.*, polypeptides encoded by biosynthetic enzyme-encoding genes), compounds (*e.g.*, precursors, substrates, intermediates or products), cofactors and the like utilized in the formation or synthesis or conversion of pyruvate to valine or isoleucine. The term “isoleucine-valine biosynthetic pathway” includes the biosynthetic pathway leading to the synthesis of valine or isoleucine in microorganisms (*e.g.*, *in vivo*) as well as the biosynthetic pathway leading to the synthesis of valine or isoleucine *in vitro*. Figure 1

includes a schematic representation of the isoleucine-valine biosynthetic pathway. Isoleucine-valine biosynthetic enzymes are depicted in bold italics and their corresponding genes indicated in italics

The term “isoleucine-valine biosynthetic enzyme” includes any enzyme
5 utilized in the formation of a compound (*e.g.*, intermediate or product) of the isoleucine-valine biosynthetic pathway. According to Figure 1, synthesis of valine from pyruvate proceeds *via* the intermediates, acetolactate, α,β -dihydroxyisovalerate (α,β -DHIV) and α -ketoisovalerate (α -KIV). Formation of acetolactate from pyruvate is catalyzed by the
10 isoleucine-valine biosynthetic enzyme acetohydroxyacid synthetase (the *ilvBN* gene product, or alternatively, the *alsS* gene product). Formation of α,β -DHIV from acetolactate is catalyzed by the isoleucine-valine biosynthetic enzyme acetohydroxyacidisomero reductase (the *ilvC* gene product). Synthesis of α -KIV from α,β -DHIV is catalyzed by the isoleucine-valine biosynthetic enzyme dihydroxyacid
15 dehydratase (the *ilvD* gene product). Moreover, valine and isoleucine can be interconverted with their respective α -keto compounds by branched chain amino acid transaminases. Based on the newly discovered HMBPA biosynthesis pathway, isoleucine-valine biosynthetic enzymes may also perform an alternative function as
enzymes in the HMBPA biosynthetic pathway described herein.

The term “3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid
20 (“HMBPA”) biosynthetic pathway” includes the alternative biosynthetic pathway involving biosynthetic enzymes and compounds (*e.g.*, substrates and the like) traditionally associated with the pantothenate biosynthetic pathway utilized in the formation or synthesis of HMBPA. The term “HMBPA biosynthetic pathway” includes the biosynthetic pathway leading to the synthesis of HMBPA in microorganisms (*e.g.*, *in vivo*) as well as the biosynthetic pathway leading to the synthesis of HMBPA *in vitro*.
25

The term “HMBPA biosynthetic enzyme” includes any enzyme utilized in the formation of a compound (*e.g.*, intermediate or product) of the HMBPA biosynthetic pathway. For example, synthesis of 2-hydroxyisovaleric acid (α -HIV) from α -ketoisovalerate (α -KIV) is catalyzed by the *panE1* or *panE2* gene product (PanE1,
30 alternatively referred to herein ketopantoate reductase or PanE2, a α -ketoacid reductase that does not significantly contribute to pantothenate biosynthesis) and/or is catalyzed by the *ilvC* gene product (alternatively referred to herein as acetohydroxyacid isomeroreductase). Formation of HMBPA from β -alanine and α -HIV is catalyzed by the *panC* gene product (alternatively referred to herein as pantothenate synthetase).

The term "3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid ("HMBPA")" includes the free acid form of HMBPA, also referred to as "3-(2-hydroxy-3-methyl-butyrylamino)-propionate" as well as any salt thereof (*e.g.*, derived by replacing the acidic hydrogen of [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid or [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionate with a cation, for example, calcium, sodium, potassium, ammonium), also referred to as a "3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid salt" or "HMBPA salt". Preferred HMBPA salts are calcium HMBPA or sodium HMBPA. HMBPA salts of the present invention include salts prepared *via* conventional methods from the free acids described herein. An HMBPA salt of the present invention can likewise be converted to a free acid form of [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid or [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionate by conventional methodology.

Various aspects of the invention are described in further detail in the following subsections.

I. Targeting Genes Encoding Various Pantothenate and/or Isoleucine-Valine(*ilv*) and/or HMBPA Biosynthetic Enzymes

In one embodiment, the present invention features targeting or modifying various biosynthetic enzymes of the pantothenate and/or isoleucine-valine(*ilv*) and/or HMBPA biosynthetic pathways. In particular, the invention features modifying various enzymatic activities associated with said pathways by modifying or altering the genes encoding said biosynthetic enzymes.

The term "gene", as used herein, includes a nucleic acid molecule (*e.g.*, a DNA molecule or segment thereof) that, in an organism, can be separated from another gene or other genes, by intergenic DNA (*i.e.*, intervening or spacer DNA which naturally flanks the gene and/or separates genes in the chromosomal DNA of the organism). Alternatively, a gene may slightly overlap another gene (*e.g.*, the 3' end of a first gene overlapping the 5' end of a second gene), said overlapping genes separated from other genes by intergenic DNA. A gene may direct synthesis of an enzyme or other protein molecule (*e.g.*, may comprise coding sequences, for example, a contiguous open reading frame (ORF) which encodes a protein) or may itself be functional in the organism. A gene in an organism, may be clustered in an operon, as defined herein, said operon being separated from other genes and/or operons by the intergenic DNA. An "isolated gene",

as used herein, includes a gene which is essentially free of sequences which naturally flank the gene in the chromosomal DNA of the organism from which the gene is derived (*i.e.*, is free of adjacent coding sequences which encode a second or distinct protein, adjacent structural sequences or the like) and optionally includes 5' and 3' regulatory sequences, for example promoter sequences and/or terminator sequences. In one embodiment, an isolated gene includes predominantly coding sequences for a protein (*e.g.*, sequences which encode *Bacillus* proteins). In another embodiment, an isolated gene includes coding sequences for a protein (*e.g.*, for a *Bacillus* protein) and adjacent 5' and/or 3' regulatory sequences from the chromosomal DNA of the organism from which the gene is derived (*e.g.*, adjacent 5' and/or 3' *Bacillus* regulatory sequences). Preferably, an isolated gene contains less than about 10 kb, 5 kb, 2 kb, 1 kb, 0.5 kb, 0.2 kb, 0.1 kb, 50 bp, 25 bp or 10 bp of nucleotide sequences that naturally flank the gene in the chromosomal DNA of the organism from which the gene is derived.

The term "operon" includes at least two adjacent genes or ORFs, optionally overlapping in sequence at either the 5' or 3' end of at least one gene or ORF. The term "operon" includes a coordinated unit of gene expression that contains a promoter and possibly a regulatory element associated with one or more adjacent genes or ORFs (*e.g.*, structural genes encoding enzymes, for example, biosynthetic enzymes). Expression of the genes (*e.g.*, structural genes) can be coordinately regulated, for example, by regulatory proteins binding to the regulatory element or by anti-termination of transcription. The genes of an operon (*e.g.*, structural genes) can be transcribed to give a single mRNA that encodes all of the proteins.

A "gene having a mutation" or "mutant gene" as used herein, includes a gene having a nucleotide sequence which includes at least one alteration (*e.g.*, substitution, insertion, deletion) such that the polypeptide or protein encoded by said mutant exhibits an activity that differs from the polypeptide or protein encoded by the wild-type nucleic acid molecule or gene. In one embodiment, a gene having a mutation or mutant gene encodes a polypeptide or protein having an increased activity as compared to the polypeptide or protein encoded by the wild-type gene, for example, when assayed under similar conditions (*e.g.*, assayed in microorganisms cultured at the same temperature). As used herein, an "increased activity" or "increased enzymatic activity" is one that is at least 5% greater than that of the polypeptide or protein encoded by the wild-type nucleic acid molecule or gene, preferably at least 5-10% greater, more preferably at least 10-25% greater and even more preferably at least 25-50%, 50-75% or

75-100% greater than that of the polypeptide or protein encoded by the wild-type nucleic acid molecule or gene. Ranges intermediate to the above-recited values, *e.g.*, 75-85%, 85-90%, 90-95%, are also intended to be encompassed by the present invention. As used herein, an "increased activity" or "increased enzymatic activity" can also include
5 an activity that is at least 1.25-fold greater than the activity of the polypeptide or protein encoded by the wild-type gene, preferably at least 1.5-fold greater, more preferably at least 2-fold greater and even more preferably at least 3-fold, 4-fold, 5-fold, 10-fold, 20-fold, 50-fold, 100-fold or greater than the activity of the polypeptide or protein encoded by the wild-type gene.

10 In another embodiment, a gene having a mutation or mutant gene encodes a polypeptide or protein having a reduced activity as compared to the polypeptide or protein encoded by the wild-type gene, for example, when assayed under similar conditions (*e.g.*, assayed in microorganisms cultured at the same temperature). A mutant gene also can encode no polypeptide or have a reduced level of production of the wild-
15 type polypeptide. As used herein, a "reduced activity" or "reduced enzymatic activity" is one that is at least 5% less than that of the polypeptide or protein encoded by the wild-type nucleic acid molecule or gene, preferably at least 5-10% less, more preferably at least 10-25% less and even more preferably at least 25-50%, 50-75% or 75-100% less than that of the polypeptide or protein encoded by the wild-type nucleic acid molecule or
20 gene. Ranges intermediate to the above-recited values, *e.g.*, 75-85%, 85-90%, 90-95%, are also intended to be encompassed by the present invention. As used herein, a "reduced activity" or "reduced enzymatic activity" can also include an activity that has been deleted or "knocked out" (*e.g.*, approximately 100% less activity than that of the polypeptide or protein encoded by the wild-type nucleic acid molecule or gene).

25 Activity can be determined according to any well accepted assay for measuring activity of a particular protein of interest. Activity can be measured or assayed directly, for example, measuring an activity of a protein isolated or purified from a cell or microorganism. Alternatively, an activity can be measured or assayed within a cell or microorganism or in an extracellular medium. For example, assaying
30 for a mutant gene (*i.e.*, said mutant encoding a reduced enzymatic activity) can be accomplished by expressing the mutated gene in a microorganism, for example, a mutant microorganism in which the enzyme is temperature-sensitive, and assaying the mutant gene for the ability to complement a temperature sensitive (Ts) mutant for enzymatic activity. A mutant gene that encodes an "increased enzymatic activity" can be one that

complements the Ts mutant more effectively than, for example, a corresponding wild-type gene. A mutant gene that encodes a "reduced enzymatic activity" is one that complements the Ts mutant less effectively than, for example, a corresponding wild-type gene.

5 It will be appreciated by the skilled artisan that even a single substitution in a nucleic acid or gene sequence (*e.g.*, a base substitution that encodes an amino acid change in the corresponding amino acid sequence) can dramatically affect the activity of an encoded polypeptide or protein as compared to the corresponding wild-type polypeptide or protein. A mutant gene (*e.g.*, encoding a mutant polypeptide or protein),
10 as defined herein, is readily distinguishable from a nucleic acid or gene encoding a protein homologue in that a mutant gene encodes a protein or polypeptide having an altered activity, optionally observable as a different or distinct phenotype in a microorganism expressing said mutant gene or producing said mutant protein or polypeptide (*i.e.*, a mutant microorganism) as compared to a corresponding
15 microorganism expressing the wild-type gene. By contrast, a protein homologue has an identical or substantially similar activity, optionally phenotypically indiscernable when produced in a microorganism, as compared to a corresponding microorganism expressing the wild-type gene. Accordingly it is not, for example, the degree of sequence identity between nucleic acid molecules, genes, protein or polypeptides that
20 serves to distinguish between homologues and mutants, rather it is the activity of the encoded protein or polypeptide that distinguishes between homologues and mutants: homologues having, for example, low (*e.g.*, 30-50% sequence identity) sequence identity yet having substantially equivalent functional activities, and mutants, for example sharing 99% sequence identity yet having dramatically different or altered functional
25 activities.

 It will also be appreciated by the skilled artisan that nucleic acid molecules, genes, protein or polypeptides for use in the instant invention can be derived from any microorganisms having a HMBPA biosynthetic pathway, an *ilv* biosynthetic pathway or a pantothenate biosynthetic pathway. Such nucleic acid molecules, genes,
30 protein or polypeptides can be identified by the skilled artisan using known techniques such as homology screening, sequence comparison and the like, and can be modified by the skilled artisan in such a way that expression or production of these nucleic acid molecules, genes, protein or polypeptides occurs in a recombinant microorganism (*e.g.*, by using appropriate promoters, ribosomal binding sites, expression or integration

vectors, modifying the sequence of the genes such that the transcription is increased (taking into account the preferable codon usage), etc., according to techniques described herein and those known in the art).

In one embodiment, the genes of the present invention are derived from a Gram positive microorganism organism (*e.g.*, a microorganism which retains basic dye, for example, crystal violet, due to the presence of a Gram-positive wall surrounding the microorganism). The term "derived from" (*e.g.*, "derived from" a Gram positive microorganism) refers to a gene which is naturally found in the microorganism (*e.g.*, is naturally found in a Gram positive microorganism). In a preferred embodiment, the genes of the present invention are derived from a microorganism belonging to a genus selected from the group consisting of *Bacillus*, *Corynebacterium* (*e.g.*, *Corynebacterium glutamicum*), *Lactobacillus*, *Lactococci* and *Streptomyces*. In a more preferred embodiment, the genes of the present invention are derived from a microorganism is of the genus *Bacillus*. In another preferred embodiment, the genes of the present invention are derived from a microorganism selected from the group consisting of *Bacillus subtilis*, *Bacillus lentimorbus*, *Bacillus lentus*, *Bacillus firmus*, *Bacillus pantothenicus*, *Bacillus amyloliquefaciens*, *Bacillus cereus*, *Bacillus circulans*, *Bacillus coagulans*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus thuringiensis*, *Bacillus halodurans*, and other Group 1 *Bacillus* species, for example, as characterized by 16S rRNA type. In another preferred embodiment, the gene is derived from *Bacillus brevis* or *Bacillus stearothermophilus*. In another preferred embodiment, the genes of the present invention are derived from a microorganism selected from the group consisting of *Bacillus licheniformis*, *Bacillus amyloliquefaciens*, *Bacillus subtilis*, and *Bacillus pumilus*. In a particularly preferred embodiment, the gene is derived from *Bacillus subtilis* (*e.g.*, is *Bacillus subtilis*-derived). The term "derived from *Bacillus subtilis*" or "*Bacillus subtilis*-derived" includes a gene which is naturally found in the microorganism *Bacillus subtilis*. Included within the scope of the present invention are *Bacillus*-derived genes (*e.g.*, *B. subtilis*-derived genes), for example, *Bacillus* or *B. subtilis* *coaX* genes, *serA* genes, *glyA* genes, *coaA* genes, *pan* genes and/or *ilv* genes.

In another embodiment, the genes of the present invention are derived from a Gram negative (excludes basic dye) microorganism. In a preferred embodiment, the genes of the present invention are derived from a microorganism belonging to a genus selected from the group consisting of *Salmonella* (*e.g.*, *Salmonella typhimurium*), *Escherichia*, *Klebsiella*, *Serratia*, and *Proteus*. In a more preferred embodiment, the

genes of the present invention are derived from a microorganism of the genus *Escherichia*. In an even more preferred embodiment, the genes of the present invention are derived from *Escherichia coli*. In another embodiment, the genes of the present invention are derived from *Saccharomyces* (e.g., *Saccharomyces cerevisiae*).

5

II. Recombinant Nucleic Acid Molecules and Vectors

The present invention further features recombinant nucleic acid molecules (e.g., recombinant DNA molecules) that include genes described herein (e.g., isolated genes), preferably *Bacillus* genes, more preferably *Bacillus subtilis* genes, even
10 more preferably *Bacillus subtilis* pantothenate biosynthetic genes and/or isoleucine-valine (*ilv*) biosynthetic genes and/or HMBPA biosynthetic genes. The term “recombinant nucleic acid molecule” includes a nucleic acid molecule (e.g., a DNA molecule) that has been altered, modified or engineered such that it differs in nucleotide sequence from the native or natural nucleic acid molecule from which the recombinant
15 nucleic acid molecule was derived (e.g., by addition, deletion or substitution of one or more nucleotides). Preferably, a recombinant nucleic acid molecule (e.g., a recombinant DNA molecule) includes an isolated gene of the present invention operably linked to regulatory sequences. The phrase “operably linked to regulatory sequence(s)” means that the nucleotide sequence of the gene of interest is linked to the regulatory
20 sequence(s) in a manner which allows for expression (e.g., enhanced, increased, constitutive, basal, attenuated, decreased or repressed expression) of the gene, preferably expression of a gene product encoded by the gene (e.g., when the recombinant nucleic acid molecule is included in a recombinant vector, as defined herein, and is introduced into a microorganism).

25 The term “regulatory sequence” includes nucleic acid sequences which affect (e.g., modulate or regulate) expression of other nucleic acid sequences (i.e., genes). In one embodiment, a regulatory sequence is included in a recombinant nucleic acid molecule in a similar or identical position and/or orientation relative to a particular gene of interest as is observed for the regulatory sequence and gene of interest as it
30 appears in nature, e.g., in a native position and/or orientation. For example, a gene of interest can be included in a recombinant nucleic acid molecule operably linked to a regulatory sequence which accompanies or is adjacent to the gene of interest in the natural organism (e.g., operably linked to “native” regulatory sequences (e.g., to the “native” promoter). Alternatively, a gene of interest can be included in a recombinant

nucleic acid molecule operably linked to a regulatory sequence which accompanies or is adjacent to another (*e.g.*, a different) gene in the natural organism. Alternatively, a gene of interest can be included in a recombinant nucleic acid molecule operably linked to a regulatory sequence from another organism. For example, regulatory sequences from
5 other microbes (*e.g.*, other bacterial regulatory sequences, bacteriophage regulatory sequences and the like) can be operably linked to a particular gene of interest.

In one embodiment, a regulatory sequence is a non-native or non-naturally-occurring sequence (*e.g.*, a sequence which has been modified, mutated, substituted, derivatized, deleted including sequences which are chemically synthesized).
10 Preferred regulatory sequences include promoters, enhancers, termination signals, anti-termination signals and other expression control elements (*e.g.*, sequences to which repressors or inducers bind and/or binding sites for transcriptional and/or translational regulatory proteins, for example, in the transcribed mRNA). Such regulatory sequences are described, for example, in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular*
15 *Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989. Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in a microorganism (*e.g.*, constitutive promoters and strong constitutive promoters), those which direct inducible expression of a nucleotide sequence in a microorganism (*e.g.*, inducible
20 promoters, for example, xylose inducible promoters) and those which attenuate or repress expression of a nucleotide sequence in a microorganism (*e.g.*, attenuation signals or repressor sequences). It is also within the scope of the present invention to regulate expression of a gene of interest by removing or deleting regulatory sequences. For example, sequences involved in the negative regulation of transcription can be removed
25 such that expression of a gene of interest is enhanced.

In one embodiment, a recombinant nucleic acid molecule of the present invention includes a nucleic acid sequence or gene that encode at least one bacterial gene product (*e.g.*, a pantothenate biosynthetic enzyme, an isoleucine-valine biosynthetic enzyme and/or a HMBPA biosynthetic enzyme) operably linked to a promoter or
30 promoter sequence. Preferred promoters of the present invention include *Bacillus* promoters and/or bacteriophage promoters (*e.g.*, bacteriophage which infect *Bacillus*). In one embodiment, a promoter is a *Bacillus* promoter, preferably a strong *Bacillus* promoter (*e.g.*, a promoter associated with a biochemical housekeeping gene in *Bacillus* or a promoter associated with a glycolytic pathway gene in *Bacillus*). In another

embodiment, a promoter is a bacteriophage promoter. In a preferred embodiment, the promoter is from the bacteriophage SP01. In a particularly preferred embodiment, a promoter is selected from the group consisting of *P*₁₅, *P*₂₆ or *P*_{veg}, having for example, the following respective sequences:

- 5 GCTATTGACGACAGCTATGGTTCACTGTCCACCAACCAAACTGTGCTCAGT
ACCGCCAATATTTCTCCCTTGAGGGGTACAAAGAGGTGTCCCTAGAAGAGAT
CCACGCTGTGTAAAAATTTTACAAAAAGGTATTGACTTTCCCTACAGGGTGT
GTAATAATTTAATTACAGGCGGGGGCAACCCCGCCTGT(SEQ ID NO:1),
GCCTACCTAGCTTCCAAGAAAGATATCCTAACAGCACAAAGAGCGGAAAGAT
10 GTTTTGTCTACATCCAGAACAACCTCTGCTAAAATTCCTGAAAAATTTTGCA
AAAAGTTGTTGACTTTATCTACAAGGTGTGGTATAATAATCTTAACAACAGC
AGGACGC (SEQ ID NO:2), and
GAGGAATCATAGAATTTTGTCAAAATAATTTTATTGACAACGTCTTATTAAC
GTTGATATAATTTAAATTTTATTGACAAAAATGGGCTCGTGTTGTACAATA
15 AATGTAGTGAGGTGGATGCAATG (SEQ ID NO:3). Additional preferred promoters
include *tef* (the translational elongation factor (TEF) promoter) and *pyc* (the pyruvate
carboxylase (PYC) promoter), which promote high level expression in *Bacillus* (*e.g.*,
Bacillus subtilis). Additional preferred promoters, for example, for use in Gram positive
microorganisms include, but are not limited to, *amy* and SPO2 promoters. Additional
20 preferred promoters, for example, for use in Gram negative microorganisms include, but
are not limited to, *cos*, *tac*, *trp*, *tet*, *trp-tet*, *lpp*, *lac*, *lpp-lac*, *lacIQ*, T7, T5, T3, *gal*, *trc*,
ara, SP6, λ -PR or λ -PL.

- In another embodiment, a recombinant nucleic acid molecule of the
present invention includes a terminator sequence or terminator sequences (*e.g.*,
25 transcription terminator sequences). The term "terminator sequences" includes
regulatory sequences that serve to terminate transcription of mRNA. Terminator
sequences (or tandem transcription terminators) can further serve to stabilize mRNA
(*e.g.*, by adding structure to mRNA), for example, against nucleases.

- In yet another embodiment, a recombinant nucleic acid molecule of the
present invention includes sequences which allow for detection of the vector containing
30 said sequences (*i.e.*, detectable and/or selectable markers), for example, genes that
encode antibiotic resistance or sequences that overcome auxotrophic mutations, for
example, *trpC*, fluorescent markers, drug markers, and/or colorimetric markers (*e.g.*,
lacZ/ β -galactosidase). In yet another embodiment, a recombinant nucleic acid molecule

of the present invention includes an artificial ribosome binding site (RBS) or a sequence that becomes transcribed into an artificial RBS. The term "artificial ribosome binding site (RBS)" includes a site within an mRNA molecule (*e.g.*, coded within DNA) to which a ribosome binds (*e.g.*, to initiate translation) which differs from a native RBS (*e.g.*, a RBS found in a naturally-occurring gene) by at least one nucleotide. Preferred artificial RBSs include about 5-6, 7-8, 9-10, 11-12, 13-14, 15-16, 17-18, 19-20, 21-22, 23-24, 25-26, 27-28, 29-30 or more nucleotides of which about 1-2, 3-4, 5-6, 7-8, 9-10, 11-12, 13-15 or more differ from the native RBS (*e.g.*, the native RBS of a gene of interest, for example, the native *panB* RBS TAAACATGAGGAGGAGAAAACATG (SEQ ID NO:4) or the native *panD* RBS ATTCGAGAAATGGAGAGAATATAATATG (SEQ ID NO:5)).

Preferably, nucleotides that differ are substituted such that they are identical to one or more nucleotides of an ideal RBS when optimally aligned for comparisons. Ideal RBSs include, but are not limited to, AGAAAGGAGGTGA (SEQ ID NO:6), TTAAGAAAGGAGGTGANNNNATG (SEQ ID NO:7), TTAGAAAGGAGGTGANNNNNATG (SEQ ID NO:8), AGAAAGGAGGTGANNNNNNATG (SEQ ID NO:9), and AGAAAGGAGGTGANNNNNNATG (SEQ ID NO:10). Artificial RBSs can be used to replace the naturally-occurring or native RBSs associated with a particular gene. Artificial RBSs preferably increase translation of a particular gene. Preferred artificial RBSs (*e.g.*, RBSs for increasing the translation of *panB*, for example, of *B. subtilis panB*) include CCCTCTAGAAGGAGGAGAAAACATG (SEQ ID NO:11) and CCCTCTAGAGGAGGAGAGAAAACATG (SEQ ID NO:12). Preferred artificial RBSs (*e.g.*, RBSs for increasing the translation of *panD*, for example, of *B. subtilis panD*) include TTAGAAAGGAGGATTTAAATATG (SEQ ID NO:13), TTAGAAAGGAGGTTTAATTAATG (SEQ ID NO:14), TTAGAAAGGAGGTGATTTAAATG (SEQ ID NO:15), TTAGAAAGGAGGTGTTTAAAATG (SEQ ID NO:16), ATTCGAGAAAGGAGGTGAATATAATATG (SEQ ID NO:17), ATTCGAGAAAGGAGGTGAATAATAATG (SEQ ID NO:18), and ATTCGTAGAAAGGAGGTGAATTAATATG (SEQ ID NO:19).

The present invention further features vectors (*e.g.*, recombinant vectors) that include nucleic acid molecules (*e.g.*, genes or recombinant nucleic acid molecules comprising said genes) as described herein. The term "recombinant vector" includes a vector (*e.g.*, plasmid, phage, phasmid, virus, cosmid or other purified nucleic acid

vector) that has been altered, modified or engineered such that it contains greater, fewer or different nucleic acid sequences than those included in the native or natural nucleic acid molecule from which the recombinant vector was derived. Preferably, the recombinant vector includes a biosynthetic enzyme-encoding gene or recombinant
5 nucleic acid molecule including said gene, operably linked to regulatory sequences, for example, promoter sequences, terminator sequences and/or artificial ribosome binding sites (RBSs), as defined herein. In another embodiment, a recombinant vector of the present invention includes sequences that enhance replication in bacteria (*e.g.*, replication-enhancing sequences). In one embodiment, replication-enhancing sequences
10 are derived from *E. coli*. In another embodiment, replication-enhancing sequences are derived from pBR322.

In yet another embodiment, a recombinant vector of the present invention includes antibiotic resistance sequences. The term "antibiotic resistance sequences" includes sequences which promote or confer resistance to antibiotics on the host
15 organism (*e.g.*, *Bacillus*). In one embodiment, the antibiotic resistance sequences are selected from the group consisting of *cat* (chloramphenicol resistance) sequences, *tet* (tetracycline resistance) sequences, *erm* (erythromycin resistance) sequences, *neo* (neomycin resistance) sequences, *kan* (kanamycin resistance) and *spec* (spectinomycin resistance) sequences. Recombinant vectors of the present invention can further include
20 homologous recombination sequences (*e.g.*, sequences designed to allow recombination of the gene of interest into the chromosome of the host organism). For example, *bpr*, *vpr*, and/or *amyE* sequences can be used as homology targets for recombination into the host chromosome. It will further be appreciated by one of skill in the art that the design of a vector can be tailored depending on such factors as the choice of microorganism to
25 be genetically engineered, the level of expression of gene product desired and the like.

IV. Recombinant Microorganisms

The present invention further features microorganisms, *i.e.*, recombinant
30 microorganisms, that include vectors or genes (*e.g.*, wild-type and/or mutated genes) as described herein. As used herein, the term "recombinant microorganism" includes a microorganism (*e.g.*, bacteria, yeast cell, fungal cell, etc.) that has been genetically altered, modified or engineered (*e.g.*, genetically engineered) such that it exhibits an altered, modified or different genotype and/or phenotype (*e.g.*, when the genetic

modification affects coding nucleic acid sequences of the microorganism) as compared to the naturally-occurring microorganism from which it was derived.

In one embodiment, a recombinant microorganism of the present invention is a Gram positive organism (e.g., a microorganism which retains basic dye, for example, crystal violet, due to the presence of a Gram-positive wall surrounding the microorganism). In a preferred embodiment, the recombinant microorganism is a microorganism belonging to a genus selected from the group consisting of *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Lactococci* and *Streptomyces*. In a more preferred embodiment, the recombinant microorganism is of the genus *Bacillus*. In another preferred embodiment, the recombinant microorganism is selected from the group consisting of *Bacillus subtilis*, *Bacillus lentimorbus*, *Bacillus lentus*, *Bacillus firmus*, *Bacillus pantothenicus*, *Bacillus amyloliquefaciens*, *Bacillus cereus*, *Bacillus circulans*, *Bacillus coagulans*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus thuringiensis*, *Bacillus halodurans*, and other Group 1 *Bacillus* species, for example, as characterized by 16S rRNA type. In another preferred embodiment, the recombinant microorganism is *Bacillus brevis* or *Bacillus stearothermophilus*. In another preferred embodiment, the recombinant microorganism is selected from the group consisting of *Bacillus licheniformis*, *Bacillus amyloliquefaciens*, *Bacillus subtilis*, and *Bacillus pumilus*.

In another embodiment, the recombinant microorganism is a Gram negative (excludes basic dye) organism. In a preferred embodiment, the recombinant microorganism is a microorganism belonging to a genus selected from the group consisting of *Salmonella*, *Escherichia*, *Klebsiella*, *Serratia*, and *Proteus*. In a more preferred embodiment, the recombinant microorganism is of the genus *Escherichia*. In an even more preferred embodiment, the recombinant microorganism is *Escherichia coli*. In another embodiment, the recombinant microorganism is *Saccharomyces* (e.g., *S. cerevisiae*).

A preferred "recombinant" microorganism of the present invention is a microorganism having a deregulated pantothenate biosynthesis pathway or enzyme, a deregulated isoleucine-valine (*ilv*) biosynthetic pathway or enzyme and/or a deregulated HMBPA biosynthetic pathway or enzyme. The term "deregulated" or "deregulation" includes the alteration or modification of at least one gene in a microorganism that encodes an enzyme in a biosynthetic pathway, such that the level or activity of the biosynthetic enzyme in the microorganism is altered or modified. Preferably, at least

one gene that encodes an enzyme in a biosynthetic pathway is altered or modified such that the gene product is enhanced or increased. The phrase "deregulated pathway" can also include a biosynthetic pathway in which more than one gene that encodes an enzyme in a biosynthetic pathway is altered or modified such that the level or activity of more than one biosynthetic enzyme is altered or modified. The ability to "deregulate" a pathway (*e.g.*, to simultaneously deregulate more than one gene in a given biosynthetic pathway) in a microorganism in some cases arises from the particular phenomenon of microorganisms in which more than one enzyme (*e.g.*, two or three biosynthetic enzymes) are encoded by genes occurring adjacent to one another on a contiguous piece of genetic material termed an "operon" (defined herein). Due to the coordinated regulation of genes included in an operon, alteration or modification of the single promoter and/or regulatory element can result in alteration or modification of the expression of each gene product encoded by the operon. Alteration or modification of the regulatory element can include, but is not limited to removing the endogenous promoter and/or regulatory element(s), adding strong promoters, inducible promoters or multiple promoters or removing regulatory sequences such that expression of the gene products is modified, modifying the chromosomal location of the operon, altering nucleic acid sequences adjacent to the operon or within the operon such as a ribosome binding site, increasing the copy number of the operon, modifying proteins (*e.g.*, regulatory proteins, suppressors, enhancers, transcriptional activators and the like) involved in transcription of the operon and/or translation of the gene products of the operon, or any other conventional means of deregulating expression of genes routine in the art (including but not limited to use of antisense nucleic acid molecules, for example, to block expression of repressor proteins). Deregulation can also involve altering the coding region of one or more genes to yield, for example, an enzyme that is feedback resistant or has a higher or lower specific activity.

In another preferred embodiment, a recombinant microorganism is designed or engineered such that at least one pantothenate biosynthetic enzyme, at least one isoleucine-valine biosynthetic enzyme, and/or at least one HMBPA biosynthetic enzyme is overexpressed. The term "overexpressed" or "overexpression" includes expression of a gene product (*e.g.*, a biosynthetic enzyme) at a level greater than that expressed prior to manipulation of the microorganism or in a comparable microorganism which has not been manipulated. In one embodiment, the microorganism can be

genetically designed or engineered to overexpress a level of gene product greater than that expressed in a comparable microorganism which has not been engineered.

Genetic engineering can include, but is not limited to, altering or modifying regulatory sequences or sites associated with expression of a particular gene (e.g., by adding strong promoters, inducible promoters or multiple promoters or by removing regulatory sequences such that expression is constitutive), modifying the chromosomal location of a particular gene, altering nucleic acid sequences adjacent to a particular gene such as a ribosome binding site, increasing the copy number of a particular gene, modifying proteins (e.g., regulatory proteins, suppressors, enhancers, transcriptional activators and the like) involved in transcription of a particular gene and/or translation of a particular gene product, or any other conventional means of deregulating expression of a particular gene routine in the art (including but not limited to use of antisense nucleic acid molecules, for example, to block expression of repressor proteins). Genetic engineering can also include deletion of a gene, for example, to block a pathway or to remove a repressor. In embodiments featuring microorganisms having deleted genes, the skilled artisan will appreciate that at least low levels of certain compounds may be required to be present in or added to the culture medium in order that the viability of the microorganism is not compromised. Often, such low levels are present in complex culture media as routinely formulated. Moreover, in processes featuring culturing microorganisms having deleted genes cultured under conditions such that commercially or industrially attractive quantities of product are produced, it may be necessary to supplement culture media with slightly increased levels of certain compounds. For example, in processes featuring culturing a microorganism having a deleted *panB* gene, at least low levels of pantothenate must be present in the media, e.g., levels such as those found in routinely formulated complex media, whereas slightly increased levels of pantothenate may be added to the media in order to produce commercially or industrially attractive amounts of, for example, HMBPA. For example, 10-30 mg/L pantothenate can be added to the media in order to produce commercially or industrially attractive amounts of HMBPA.

In another embodiment, the microorganism can be physically or environmentally manipulated to overexpress a level of gene product greater than that expressed prior to manipulation of the microorganism or in a comparable microorganism which has not been manipulated. For example, a microorganism can be treated with or cultured in the presence of an agent known or suspected to increase transcription of a

particular gene and/or translation of a particular gene product such that transcription and/or translation are enhanced or increased. Alternatively, a microorganism can be cultured at a temperature selected to increase transcription of a particular gene and/or translation of a particular gene product such that transcription and/or translation are enhanced or increased.

V. *Culturing and Fermenting Recombinant Microorganisms*

The term "culturing" includes maintaining and/or growing a living microorganism of the present invention (*e.g.*, maintaining and/or growing a culture or strain). In one embodiment, a microorganism of the invention is cultured in liquid media. In another embodiment, a microorganism of the invention is cultured in solid media or semi-solid media. In a preferred embodiment, a microorganism of the invention is cultured in media (*e.g.*, a sterile, liquid media) comprising nutrients essential or beneficial to the maintenance and/or growth of the microorganism (*e.g.*, carbon sources or carbon substrate, for example carbohydrate, hydrocarbons, oils, fats, fatty acids, organic acids, and alcohols; nitrogen sources, for example, peptone, yeast extracts, meat extracts, malt extracts, urea, ammonium sulfate, ammonium chloride, ammonium nitrate and ammonium phosphate; phosphorus sources, for example, phosphoric acid, sodium and potassium salts thereof; trace elements, for example, magnesium, iron, manganese, calcium, copper, zinc, boron, molybdenum, and/or cobalt salts; as well as growth factors such as amino acids, vitamins, growth promoters and the like).

Preferably, microorganisms of the present invention are cultured under controlled pH. The term "controlled pH" includes any pH which results in production of the desired product (*e.g.*, HMBPA). In one embodiment microorganisms are cultured at a pH of about 7. In another embodiment, microorganisms are cultured at a pH of between 6.0 and 8.5. The desired pH may be maintained by any number of methods known to those skilled in the art.

Also preferably, microorganisms of the present invention are cultured under controlled aeration. The term "controlled aeration" includes sufficient aeration (*e.g.*, oxygen) to result in production of the desired product (*e.g.*, HMBPA). In one embodiment, aeration is controlled by regulating oxygen levels in the culture, for example, by regulating the amount of oxygen dissolved in culture media. Preferably, aeration of the culture is controlled by agitating the culture. Agitation may be provided by a propeller or similar mechanical agitation equipment, by revolving or shaking the

culture vessel (*e.g.*, tube or flask) or by various pumping equipment. Aeration may be further controlled by the passage of sterile air or oxygen through the medium (*e.g.*, through the fermentation mixture). Also preferably, microorganisms of the present invention are cultured without excess foaming (*e.g.*, *via* addition of antifoaming agents).

- 5 Moreover, microorganisms of the present invention can be cultured under controlled temperatures. The term "controlled temperature" includes any temperature which results in production of the desired product (*e.g.*, HMBPA). In one embodiment, controlled temperatures include temperatures between 15°C and 95°C. In another embodiment, controlled temperatures include temperatures between 15°C and 70°C.
- 10 Preferred temperatures are between 20°C and 55°C, more preferably between 30°C and 50°C.

- Microorganisms can be cultured (*e.g.*, maintained and/or grown) in liquid media and preferably are cultured, either continuously or intermittently, by conventional culturing methods such as standing culture, test tube culture, shaking culture (*e.g.*, rotary
- 15 shaking culture, shake flask culture, etc.), aeration spinner culture, or fermentation. In a preferred embodiment, the microorganisms are cultured in shake flasks. In a more preferred embodiment, the microorganisms are cultured in a fermentor (*e.g.*, a fermentation process). Fermentation processes of the present invention include, but are not limited to, batch, fed-batch and continuous processes or methods of fermentation.
- 20 The phrase "batch process" or "batch fermentation" refers to a closed system in which the composition of media, nutrients, supplemental additives and the like is set at the beginning of the fermentation and not subject to alteration during the fermentation, however, attempts may be made to control such factors as pH and oxygen concentration to prevent excess media acidification and/or microorganism death. The phrase "fed-
- 25 batch process" or "fed-batch" fermentation refers to a batch fermentation with the exception that one or more substrates or supplements are added (*e.g.*, added in increments or continuously) as the fermentation progresses. The phrase "continuous process" or "continuous fermentation" refers to a system in which a defined fermentation media is added continuously to a fermentor and an equal amount of used or
- 30 "conditioned" media is simultaneously removed, preferably for recovery of the desired product (*e.g.*, HMBPA). A variety of such processes have been developed and are well-known in the art.

The phrase "culturing under conditions such that a desired compound is produced" includes maintaining and/or growing microorganisms under conditions (*e.g.*, temperature, pressure, pH, duration, etc.) appropriate or sufficient to obtain production of the desired compound or to obtain desired yields of the particular compound being produced. For example, culturing is continued for a time sufficient to produce the desired amount of a compound (*e.g.*, HMBPA). Preferably, culturing is continued for a time sufficient to substantially reach suitable production of the compound (*e.g.*, a time sufficient to reach a suitable concentration of HMBPA or suitable ratio of HMBPA:pantothenate). In one embodiment, culturing is continued for about 12 to 24 hours. In another embodiment, culturing is continued for about 24 to 36 hours, 36 to 48 hours, 48 to 72 hours, 72 to 96 hours, 96 to 120 hours, 120 to 144 hours, or greater than 144 hours. In yet another embodiment, microorganisms are cultured under conditions such that at least about 5 to 10 g/L of compound are produced in about 36 hours, at least about 10 to 20 g/L compound are produced in about 48 hours, or at least about 20 to 30 g/L compound in about 72 hours. In yet another embodiment, microorganisms are cultured under conditions such that at least a ratio of HMBPA:HMBPA+pantothenate of 1:10 is achieved (*i.e.*, 10% HMBPA versus 90% pantothenate, for example, as determined by comparing the peaks when a sample of product is analyzed by HPLC), preferably such that at least a ratio of 2:10 is achieved (20% HMBPA versus 90% pantotheante), more preferably such that a ratio of at least 2.5:10 is achieved (25% HMBPA versus 75% pantotheante), more preferably at least 3:10 (30% HMBPA versus 70% pantotheante), 4:10 (40% HMBPA versus 60% pantotheante), 5:10 (50% HMBPA versus 50% pantotheante), 6:10 (60% HMBPA versus 40% pantotheante), 7:10 (70% HMBPA versus 30% pantotheante), 8:10 (80% HMBPA versus 20% pantotheante), 9:10 (90% HMBPA versus 10% pantotheante) or greater.

The methodology of the present invention can further include a step of recovering a desired compound (*e.g.*, HMBPA). The term "recovering" a desired compound includes extracting, harvesting, isolating or purifying the compound from culture media. Recovering the compound can be performed according to any conventional isolation or purification methodology known in the art including, but not limited to, treatment with a conventional resin (*e.g.*, anion or cation exchange resin, non-ionic adsorption resin, etc.), treatment with a conventional adsorbent (*e.g.*, activated charcoal, silicic acid, silica gel, cellulose, alumina, etc.), alteration of pH, solvent extraction (*e.g.*, with a conventional solvent such as an alcohol, ethyl acetate, hexane and

the like), dialysis, filtration, concentration, crystallization, recrystallization, pH adjustment, lyophilization and the like. For example, a compound can be recovered from culture media by first removing the microorganisms from the culture. Media are then passed through or over a cation exchange resin to remove cations and then through
5 or over an anion exchange resin to remove inorganic anions and organic acids having stronger acidities than the compound of interest. The resulting compound can subsequently be converted to a salt (*e.g.*, a calcium salt) as described herein.

Preferably, a desired compound of the present invention is “extracted”, “isolated” or “purified” such that the resulting preparation is substantially free of other
10 media components (*e.g.*, free of media components and/or fermentation byproducts). The language “substantially free of other media components” includes preparations of the desired compound in which the compound is separated from media components or fermentation byproducts of the culture from which it is produced. In one embodiment, the preparation has greater than about 80% (by dry weight) of the desired compound
15 (*e.g.*, less than about 20% of other media components or fermentation byproducts), more preferably greater than about 90% of the desired compound (*e.g.*, less than about 10% of other media components or fermentation byproducts), still more preferably greater than about 95% of the desired compound (*e.g.*, less than about 5% of other media components or fermentation byproducts), and most preferably greater than about 98-99% desired
20 compound (*e.g.*, less than about 1-2% other media components or fermentation byproducts). When the desired compound has been derivatized to a salt, the compound is preferably further free of chemical contaminants associated with the formation of the salt. When the desired compound has been derivatized to an alcohol, the compound is
25 preferably further free of chemical contaminants associated with the formation of the alcohol.

In an alternative embodiment, the desired compound is not purified from the microorganism, for example, when the microorganism is biologically non-hazardous (*e.g.*, safe). For example, the entire culture (or culture supernatant) can be used as a source of product (*e.g.*, crude product). In one embodiment, the culture (or culture
30 supernatant) is used without modification. In another embodiment, the culture (or culture supernatant) is concentrated. In yet another embodiment, the culture (or culture supernatant) is dried or lyophilized.

Preferably, a production method of the present invention results in production of the desired compound at a significantly high yield. The phrase “significantly high yield” includes a level of production or yield which is sufficiently elevated or above what is usual for comparable production methods, for example, which is elevated to a level sufficient for commercial production of the desired product (*e.g.*,
5 production of the product at a commercially feasible cost). In one embodiment, the invention features a production method that includes culturing a recombinant microorganism under conditions such that the desired product (*e.g.*, HMBPA) is produced at a level greater than 2 g/L. In another embodiment, the invention features a
10 production method that includes culturing a recombinant microorganism under conditions such that the desired product (*e.g.*, HMBPA) is produced at a level greater than 10 g/L. In another embodiment, the invention features a production method that includes culturing a recombinant microorganism under conditions such that the desired product (*e.g.*, HMBPA) is produced at a level greater than 20 g/L. In yet another
15 embodiment, the invention features a production method that includes culturing a recombinant microorganism under conditions such that the desired product (*e.g.*, HMBPA) is produced at a level greater than 30 g/L. In yet another embodiment, the invention features a production method that includes culturing a recombinant microorganism under conditions such that the desired product (*e.g.*, HMBPA) is
20 produced at a level greater than 40 g/L. The invention further features a production method for producing the desired compound that involves culturing a recombinant microorganism under conditions such that a sufficiently elevated level of compound is produced within a commercially desirable period of time.

Depending on the biosynthetic enzyme or combination of biosynthetic
25 enzymes manipulated, it may be desirable or necessary to provide (*e.g.*, feed) microorganisms of the present invention at least one biosynthetic precursor such that the desired compound or compounds are produced. The term “biosynthetic precursor” or “precursor” includes an agent or compound which, when provided to, brought into contact with, or included in the culture medium of a microorganism, serves to enhance
30 or increase biosynthesis of the desired product. In one embodiment, the biosynthetic precursor or precursor is aspartate. In another embodiment, the biosynthetic precursor or precursor is β -alanine. The amount of aspartate or β -alanine added is preferably an amount that results in a concentration in the culture medium sufficient to enhance productivity of the microorganism (*e.g.*, a concentration sufficient to enhance production

of HMBPA. The term "excess β -alanine" includes β -alanine levels increased or higher than those routinely utilized for culturing the microorganism in question. For example, culturing the *Bacillus* microorganisms described in the instant Examples is routinely done in the presence of about 0-5 g/L β -alanine. Accordingly, excess β -alanine levels
5 can include levels of about 5-10 g/L or more preferably about 5-20 g/L β -alanine. Biosynthetic precursors of the present invention can be added in the form of a concentrated solution or suspension (e.g., in a suitable solvent such as water or buffer) or in the form of a solid (e.g., in the form of a powder). Moreover, biosynthetic precursors of the present invention can be added as a single aliquot, continuously or intermittently
10 over a given period of time.

In yet another embodiment, the biosynthetic precursor is valine. In yet another embodiment, the biosynthetic precursor is α -ketoisovalerate. Preferably, valine or α -ketoisovalerate is added in an amount that results in a concentration in the medium sufficient for production of the desired product (e.g., HMBPA) to occur. The term
15 "excess α -KIV" includes α -KIV levels increased or higher than those routinely utilized for culturing the microorganism in question. For example, culturing the *Bacillus* microorganisms described in the instant Examples can be done in the presence of about 0-5 g/L α -KIV. Accordingly, excess α -KIV levels can include levels of about 5-10 g/L, and more preferably about 5-20 g/L. The term "excess valine" includes valine levels
20 increased or higher than those routinely utilized for culturing the microorganism in question. For example, culturing the *Bacillus* microorganisms described in the instant Examples is routinely done in the presence of about 0-0.5 g/L valine. Accordingly, excess valine levels can include levels of about 0.5-5 g/L, preferably about 5-20 g/L valine. Biosynthetic precursors are also referred to herein as "supplemental biosynthetic
25 substrates".

Moreover, certain aspects of the present invention include culturing microorganisms (e.g., recombinant microorganisms) under conditions of increased steady state glucose, decreased steady state dissolved oxygen and/or decreased serine. The term "increased steady state glucose" includes steady state glucose levels increased
30 or higher than those routinely utilized for culturing the microorganism in question. For example, culturing the *Bacillus* microorganisms described in the instant Examples is routinely done in the presence of about 0.2-1.0 g/L steady state glucose. Accordingly, increased steady state glucose levels can include levels of about 1-2 g/l, about 2-5 g/l, and preferably about 5-20 g/L steady state glucose. The term "decreased steady state

dissolved oxygen” includes steady state dissolved oxygen levels less or lower than those routinely utilized for culturing the microorganism in question and, for example, inversely correlates with increased steady state glucose levels. For example, culturing the *Bacillus* microorganisms described in the instant Examples is routinely done in the presence of about 10-30% dissolved oxygen. Accordingly, decreased steady state dissolved oxygen can include levels of about 0-10%, and preferably about 0-5% steady state dissolved oxygen. The term “reduced serine” includes serine levels within the lower range of those routinely utilized for culturing the microorganism in question. For example, culturing the *Bacillus* microorganisms described in the instant Examples is routinely done in the presence of about 0-0.5 g/L serine. Accordingly, reduced serine levels can include, for example, levels of 0-0.1 g/L serine.

Another aspect of the present invention includes biotransformation processes which feature the recombinant microorganisms described herein. The term “biotransformation process”, also referred to herein as “bioconversion processes”, includes biological processes which results in the production (*e.g.*, transformation or conversion) of appropriate substrates and/or intermediate compounds into a desired product (*e.g.*, HMBPA).

In one embodiment, the invention features a biotransformation process for the production of HMBPA comprising contacting a microorganism which overexpresses a reductase (*e.g.*, overexpresses PanE1, PanE2 and/or IlvC) with appropriate substrates or precursors under conditions such that HMBPA is produced and recovering said HMBPA. In another embodiment, the invention features a biotransformation process for the production of HMBPA comprising contacting a microorganism which has a reduced or deleted PanB activity with appropriate substrates or precursors under conditions such that HMBPA is produced and recovering said HMBPA. In yet another embodiment, the invention features a biotransformation process for the production of HMBPA comprising contacting a microorganism which overexpresses at least one reductase and has a reduced or deleted PanB activity with appropriate substrates or precursors under conditions such that HMBPA is produced and recovering said HMBPA. Preferred recombinant microorganisms for carrying out the above-described biotransformations include the recombinant microorganisms described herein. In yet another embodiment, the invention features a biotransformation reaction that includes contacting α -HIV and β -alanine with isolated or purified PanC under conditions such that HMBPA is produced. α -HIV can optionally be obtained by

contacting α -KIV with purified or isolated reductase (*e.g.*, PanE1, PanE2 and/or IlvC) and a source of reducing equivalent, for example, NADH. Conditions under which α -HIV or HMBPA are produced can include any conditions which result in the desired production of α -HIV or HMBPA, respectively. In yet another embodiment, the present invention includes a method of producing α -HIV that includes culturing a microorganism that overexpresses PanE1 and/or PanE2, and/or IlvC and has a reduced or deleted PanC or PanD (to reduce HMBPA or β -alanine synthesis, respectively) under conditions such that α -HIV is produced.

The microorganism(s) and/or enzymes used in the biotransformation reactions are in a form allowing them to perform their intended function (*e.g.*, producing a desired compound). The microorganisms can be whole cells, or can be only those portions of the cells necessary to obtain the desired end result. The microorganisms can be suspended (*e.g.*, in an appropriate solution such as buffered solutions or media), rinsed (*e.g.*, rinsed free of media from culturing the microorganism), acetone-dried, immobilized (*e.g.*, with polyacrylamide gel or k-carrageenan or on synthetic supports, for example, beads, matrices and the like), fixed, cross-linked or permeablized (*e.g.*, have permeablized membranes and/or walls such that compounds, for example, substrates, intermediates or products can more easily pass through said membrane or wall).

20

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application are incorporated herein by reference.

25

EXAMPLES

Example I: Discovery and Characterization of the [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA) Biosynthetic Pathway

5 In developing *Bacillus* strains for the production of pantothenate, various genetic manipulations were made to enzymes involved in the pantothenate biosynthetic pathway and the isoleucine-valine (*ilv*) pathway (Figure 1) as described in U.S. Patent Application Serial No. 09/400,494 and U.S. Patent Application Serial No. 09/667,569. For example, strains having a deregulated *panBCD* operon and/or having deregulated
10 *panE1* exhibited enhanced pantothenate production (when cultured in the presence of β -alanine and α -KIV). Strains further deregulated for *ilvBNC* and *ilvD* exhibited enhanced pantothenate production in the presence of only β -alanine. Moreover, it was possible to achieve β -alanine independence by further deregulating *panD*.

An exemplary strain is PA824, a tryptophan prototroph, Spec and Tet
15 resistant, deregulated for *panBCD* at the *panBCD* locus, deregulated for *panE1* at the *panE1* locus (two genes in the *B. subtilis* genome are homologous to *E. coli panE*, *panE1* and *panE2*, the former encoding the major ketopantoate reductase involved in pantothenate production, while *panE2* does not contribute to pantothenate synthesis (U.S. Patent Application Serial No. 09/400,494), deregulated for *ilvD* at the *ilvD* locus,
20 overexpressing an *ilvBNC* cassette at the *amyE* locus, and overexpressing *panD* at the *bpr* locus.

The production of pantothenic acid by PA824 was investigated in 14 L fermentor vessels. The composition of the batch and feed media are as follows.

25

BATCH

	MATERIAL	g/L (final)
1	Yeast extract	10
2	Na Glutamate	5
3	(NH ₄) ₂ SO ₄	8
4	KH ₂ PO ₄	5
5	K ₂ HPO ₄	7.6

Added After Sterilization and Cool Down

1	Glucose	2.5
2	CaCl ₂	0.1
3	MgCl ₂	1
4	Sodium Citrate	1
5	FeSO ₄ ·7 H ₂ O	0.01
5	SM-1000X	1 ml

The final volume of the batch medium is 6 L. The trace element solution SM-1000X has following composition: 0.15 g Na₂MoO₄·2 H₂O, 2.5 g H₃BO₃, 0.7 g CoCl₂·6 H₂O, 0.25 g CuSO₄·5 H₂O, 1.6 g MnCl₂·4 H₂O, 0.3 g ZnSO₄·7 H₂O are dissolved in water (final volume 1L).

The batch medium was inoculated with 60 ml of shake flask PA824 culture (OD=10 in SVY medium: Difco Veal Infusion broth 25 g, Difco Yeast extract 5 g, Sodium Glutamate 5 g, (NH₄)₂SO₄ 2.7 g in 740 ml H₂O, autoclave; add 200 ml sterile 1 M K₂HPO₄ (pH 7) and 60 ml sterile 50% Glucose solution (final volume 1L)). The fermentation was run at 43 °C at an air flow rate of 12 L/min as a glucose limited fed batch. The initial batched glucose (2.5 g/L) was consumed during exponential growth. Afterwards glucose concentrations were maintained between 0.2-1 g/L by continuous feeding of FEED solution as follows.

FEED

	MATERIAL	g/L (final)
1	Glucose	550
2	CaCl ₂	0.1
3	SM-1000X	3 ml

The variable feed rate pump was computer controlled and linked to the glucose concentration in the tank by an algorithm. In this example the total feeding was 6 L.

During fermentation the pH was set at 7.2. Control was achieved by pH measurements linked to computer control. The pH value was maintained by feeding either a 25% NH₃-solution or a 20% H₃PO₄-solution. NH₃ acts simultaneously as a N-

source for the fermentation. The dissolved oxygen concentration $[pO_2]$ was set at 30% by regulation of the agitation and aeration rate. Foaming was controlled by addition of silicone oil. After the stop of the addition of the feed solution, in this example after 48 h, the fermentation was continued until the $[pO_2]$ value reached 95%. Then the fermentation was stopped by killing the microorganism through sterilization for 30 min. The successful sterilization was proven by plating a sample of the fermentation broth on agar plates. The pantothenate titer in the fermentation broth was 21.7 g/L after sterilization and removal of the cells by centrifugation (determined by HPLC analysis).

For HPLC analysis the fermentation broth sample was diluted with sterile water (1:40). 5 μ l of this dilution was injected into a HPLC column (Aqua C18, 5 μ m, 150*2.0 mm, PhenomenexTM). Temperature of the column was held at 40°C. Mobile phase A was 14.8 mM H_3PO_3 , mobile phase B 100% Acetonitrile. Flow rate was constant at 0.5 mL/min. A gradient was applied :

15	start: 2% mobile phase B 0-3 min linear increase to 3% mobile phase B 3-3.5 min linear increase to 20% mobile phase B
----	---

The detection was carried out by an UV-detector (210 nm). Run time was 7 min with an additional 3 min posttime. The retention time for pantothenic acid is 3.9 minutes. The HPLC chromatogram for the above mentioned sample is given in Figure 4.

Identification of compound related to peak with retention time 4.7 minutes

Under the described fermentation conditions, PA824 routinely yields approximately 20-30 g/L pantothenate. In addition to producing significant quantities of pantothenate, it was discovered a second compound eluted with an approximate retention time of 4.7 minutes in this system. The second prominent product formed in the fermentation was shown to be [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA) (also referred to herein as “ β -alanine 2-(R)-hydroxyisovalerate”, “ β -alanine 2-hydroxyisovalerate”, and/or “ β -alanyl- α -hydroxyisovalerate”). It was identified by its mass spectrum (Figure 5; relative monoisotopic mass 189), ¹H- and ¹³C-NMR (data not shown) after chromatographic purification by reverse phase flash chromatography (mobile phase 10 mM KH_2PO_4 , with increasing contents of acetonitrile (1-50%)).

In order to verify the identity of the compound, deliberate synthesis of racemic β -alanine 2-hydroxyisovalerate was performed as follows. β -alanine (2.73 g / 30 mmol) and sodium methoxide (5.67 g of a 30% solution in methanol / 31.5 mmol) were dissolved in methanol (40 mL). Methyl 2-hydroxyisovalerate (2-hydroxy-3-methylbutyric acid methyl ester) (3.96g / 30 mmol) was added and refluxed for 18 hours. Methanol was then removed by rotavap and replaced by tert-butanol (50 mL). Potassium tert-butoxide was added (50 mg) and refluxed for 26 hours. The solvent was removed *in vacuo*, the residue dissolved in water (50 mL) and passed through a strongly acidic ion-exchange resin (H⁺-form Lewatite™ S 100 G1; 100 mL). More water is used to rinse the ion exchanger. The aqueous eluates are combined and the water removed *in vacuo*. The residue is subjected to flash chromatography (silica gel; 2% acetic acid in ethyl acetate as eluent) and the product fractions evaporated to give a solid residue. The residue was recrystallized from ethyl acetate / toluene (10 mL / 20 mL, respectively) and analytically pure HMBPA (β -alanine 2-hydroxyisovalerate) was obtained, which showed a relative monoisotopic mass of 190 (189 + H⁺) in the mass spec and the same ¹H-NMR resonances as the product obtained from fermentation.

The biosynthetic pathway resulting in HMBPA production is set forth in Figure 2. The chemical structure of [R]-3-(2-hydroxy-3-methyl-butrylamino)-propionic acid (HMBPA) is depicted in Figure 3. As depicted in Figure 2, HMBPA is the condensation product of α -hydroxyisovaleric acid (α -HIV) and β -alanine, catalyzed by the PanC enzyme. α -HIV is generated by reduction of α -KIV, a reaction which is catalyzed by the reductases PanE (*e.g.*, PanE1 and/or PanE2) and/or IlvC.

Based on the chemical structure and biosynthetic pathway leading to HMBPA production, the present inventors formulated the following model to describe the interaction between the previously known pantothenate and isoleucine-valine (*ilv*) pathways and the newly characterized HMBPA biosynthetic pathway. In at least one aspect, the model states that there exist at least two pathways in microorganisms that compete for α -KIV, the substrate for the biosynthetic enzyme PanB, namely the pantothenate biosynthetic pathway and the HMBPA biosynthetic pathway. (A third and fourth pathway competing for α -KIV are those resulting in the production of valine or leucine from α -KIV, see *e.g.*, Figure 1). At least the pantothenate biosynthetic pathway and the HMBPA biosynthetic pathway further produce competitive substrates for the enzyme PanC, namely α -HIV and pantoate. The model predicts that reducing PanB activity will increase α -KIV availability for α -HIV synthesis (and ultimately, HMBPA

synthesis) and decrease the amount of pantoate and/or pantothenate synthesized by a microorganism. Conversely, increasing PanB activity will increase pantoate and ketopantoate availability for pantoate/pantothenate synthesis. The following examples provide experimental support for the model and further exemplify processes for increasing the production of HMBPA based on the model.

EXAMPLES II-VI:

For Examples II-VI, quantitation of pantothenate and/or HMBPA was performed as follows. Aliquots of fermentation media were diluted 1:100 and aliquots of test tube cultures were diluted 1:10 in water or 5% acetonitrile prior to injection on a Phenomenex Aqua™ 5 μ C18 HPLC column (250 x 4.60mm, 125A). Mobile phases were A = 5% acetonitrile, 50 mM monosodium phosphate buffer adjusted to pH 2.5 with phosphoric acid; and B = 95% acetonitrile, 5% H₂O.

Linear gradients were as follows.

Minutes	Solvent A	Solvent B
0	100%	0%
16	100%	0%
17	0%	100%
20	0%	100%
21	100%	0%

Additional parameters and apparatus were as follows: Flow rate = 1.0 ml/min; Injection volume = 20 μ l; Detector = Hewlett Packard 1090 series DAD UV detector-3014, Signal A = 197 nm, ref. = 450 nm, Firmware revision E; Column heater = Oven temperature 40°C; Hardware = Hewlett Packard Kayak™ XA; and Software = Hewlett Packard Chemstation Plus™ family revision A.06.03[509].

Under these fermentation conditions, PA824 routinely yields approximately 30-40 g/L pantothenate. HMBPA elutes at approximately 13 minutes in this system.

Example II: Ketopantoate Reductase Contributes to the Production of HMBPA and Increasing Ketopantoate Reductase Activity in *Bacillus* Results in Enhanced HMBPA Production

5 As described in Example I, a novel HPLC peak corresponding to HMBPA was observed in microorganisms overexpressing *panE1* indicating that increased ketopantoate reductase contributes to the production of HMBPA (in addition to production of pantothenate). As mentioned previously, two genes in the *B. subtilis* genome are homologous to the *E. coli panE* gene encoding ketopantoate reductase and
10 have been named *panE1* and *panE2*. In *Bacillus*, the *panE1* gene encodes the major ketopantoate reductase involved in pantothenate production, while *panE2* does not contribute to pantothenate synthesis. In fact, overexpression of *panE2* from a *P₂₆* promoter leads to a reduction in pantothenate titer (see e.g., U.S. Patent Application Serial No. 09/400,494).

15 Accordingly, it was tested whether, beside being produced by the *panE1* gene product, it was possible that a significant portion of the α -HIV necessary to make HMBPA was being produced by the *panE2* gene product. It was hypothesized that the *panE2* gene product is an enzyme that can reduce α -KIV to α -HIV, but that can not significantly reduce ketopantoate to pantoate.

20 To test the hypothesis, *panE2* was deleted from pantothenate production strain PA824 (described in Example I) by transforming with a Δ *panE2::cat* cassette from chromosomal DNA of strain PA248 (Δ *panE2::cat*) (set forth as SEQ ID NO:24, for construction see e.g., U.S. Patent Application Serial No. 09/400,494) to give strain PA919. Three isolates of PA919 were compared to PA824 for pantothenate and
25 HMBPA production in test tube cultures grown in SVY plus β -alanine.

Table 1. Production of pantothenate and HMBPA by derivatives of PA824 and PA880 grown at 43°C in 48 hour test tube cultures of SVY glucose + β -alanine⁵.

Strain	new trait	parent	OD ₆₀₀	5	
				[pan] g/l	[HMBPA] g/l
PA824	—		13.9	4.3	0.64
PA919-1	Δ panE2::cat	PA824	13.2	4.2	0.15
PA919-2	"	"	14.8	3.8	0.13
PA919-3	"	"	18.0	5.5	0.14

As indicated by the data in Table 1, all three isolates of PA919 produced about four-fold lower HMBPA than PA824 demonstrating that the *panE2* gene product is a potent contributor to HMBPA synthesis. Moreover, significant increases in HMBPA production can be achieved simply by overexpression of *panE2*. An exemplary plasmid for the overexpression of *panE2*, named pAN238, is set forth as SEQ ID NO:25 (Figure 10).

10 Example III. Increasing Production of HMBPA by Reducing PanB Activity in Microorganisms.

Strains derived from PA365 (the RL-1 lineage equivalent of PA377, described in U.S. Patent Application Serial No. 09/667,569) which are deleted for the *P₂₆ panBCD* cassette and which contain a *P₂₆ panC*D* cassette amplified at the *vpr* locus and either the wild type *P₂₆ panB* cassette (PA666) or a *P₂₆ ΔpanB* cassette (PA664) amplified at the *bpr* locus were constructed as follows. An alignment of the C-terminal amino acids of known or suspected PanB proteins is shown in Figure 6. Three regions called 1, 2 and 3, that were identified having conserved or semi-conserved amino acid residues, are indicated by arrows at the top of the figure. The *B. subtilis* PanB protein (RBS02239) is underlined. Two of the PanB proteins (RCY14036 and CAB56202.1) are missing region 3 while the latter PanB protein is also missing region 2 and has non-conserved amino acid residues occupying region 1.

B. subtilis PanB variants were created that were missing regions 1, 2 and 3. The desired variants were created by designing 3' PCR primers to amplify the *B. subtilis pan B* gene such that region 3, regions 2 and 3, or all three regions would be missing from the final product. The PCR products were generated and cloned into *E. coli* expression vector pASK-1BA3, creating plasmids pAN446, pAN447, and pAN448, respectively. The plasmids were then transformed into *E. coli* strain SJ2 that contains the *panB6* mutation to test for complementation. Only pAN446, which is missing region 3, was able to complement. This indicates that region 3 is not essential for *B. subtilis* PanB activity but that region 2 is required for activity or stability.

The next step in this analysis was to transfer the *panB* gene from pAN446 to a *B. subtilis* expression vector and then introduce it into a strain appropriate for testing activity of the encoded PanB protein in *B. subtilis*. To do this, a strain that is deleted for the *P*₂₆ *panBCD* operon was first created. This was accomplished by first
 5 inserting a *cat* gene between the *Bse*RI site located just upstream of the *panB* RBS and the *Bg*/II site located in *panD*, creating plasmid pAN624 (*Figure 7*). The sequence of pAN624 is set forth as SEQ ID NO:20. The resulting deletion-substitution mutation (Δ *panBCD::cat624*), which removes all of *panB* and *panC*, was crossed into PA354 by transformation, with selection for resistance to chloramphenicol on plates supplemented
 10 with 1 mM pantothenate. One of the transformants was saved and named PA644. Chromosomal DNA isolated from PA644 was analyzed by PCR and was shown to contain the deletion-substitution mutation. As expected, PA644 requires pantothenate for growth but retains the engineered *ilv* genes (*P*₂₆*ilvBNC* *P*₂₆*ilvD*) as well as the *P*₂₆*panE1* gene originally present in PA354. Thus, it has all the enzymes involved in pantoate
 15 synthesis overproduced except PanB. The gene containing the shortest *panB* deletion was inserted into *B. subtilis* expression vector pOTP61 (described in US patent application Serial No. 09/667,569), creating plasmid pAN627. At the same time, a wild-type *panB* control gene was inserted into pOTP61, creating plasmid pAN630. The *Not*I fragments of each plasmid, lacking *E. coli* vector sequences, were ligated and
 20 transformed into PA644, with selection for resistance to tetracycline.

One transformant from each transformation was saved and further transformed with chromosomal DNA from PA628 with selection for Pan⁺. PA628 contains a multicopy *P*₂₆*panC*D* expression plasmid (pAN620) integrated at the *vpr* locus. In order to determine the effects of the *panB* gene mutation directly on
 25 pantothenate production, plasmid pAN620, set forth as SEQ ID NO:21 and illustrated schematically in *Figure 8*, provides the remaining two enzymes required for pantothenate synthesis (PanC and PanD). Four transformants from each transformation were isolated, grown in SVY medium containing 10 g/L aspartate for 48 hours, then assayed for pantothenate production. Transformants with the 3' deleted *panB* gene were
 30 named PA664 and those containing the wild-type gene were called PA666. The data showed that the 3' deleted *panB* gene in PA664 encodes a PanB protein with greatly reduced activity. To test for HMBPA production, test tube cultures of PA365, PA666, and PA664 were grown in SVY + aspartate medium with and without added α -KIV or

pantoate for 48 hours and then assayed for HMBPA and pantothenate as described previously.

Table 2. Effect of PanB activity and addition of precursors on HMBPA and pantothenate production, 48 hour test tube culture data, SVY + aspartate (10 g/L) medium.

Strain	pan operon		panC*D plasmid		panB plasmid		no additions		+ α -KIV (5 g/L)		+ pantoate (5 g/L)	
PA365	P_{26}	panBCD	NONE	NONE	NONE		[pan] (g/L)	HMBPA peak*	[pan] (g/L)	HMBPA peak	[pan] (g/L)	HMBPA peak
PA666	Δ panBCD::cat		pan620	pan630			3.0	0.71	3.2	1.28	4.8	0.38
PA664	Δ panBCD::cat		pan620	pan627			3.7	0.55	3.3	1.70	5.2	0.26
							0.3	1.39	0.6	1.76	2.5	0.74

* HMBPA peak = peak area $\times 10^{-3}$

The data presented in Table 2 demonstrate that in the absence of supplements, PA664 produced the most HMBPA while PA666 produced the least, indicating an inverse correlation between PanB activity and HMBPA production. This is consistent with the model which predicts that the two pathways compete for α -KIV, the substrate for PanB, and produce competitive substrates for PanC; lowering PanB activity would be expected to increase α -KIV availability for α -HIV synthesis and increase HMBPA production, correspondingly decreasing the amount of pantoate synthesized. When α -KIV is added to the medium, all three strains produced significantly more HMBPA. This result evidences that α -KIV is a precursor to HMBPA, as described in Figure 2, and that excess α -KIV favors HMBPA production. This result also suggests that synthesis of HMBPA is at least partially due to an overflow effect of excess α -KIV production. When pantoate was added to the medium, HMBPA was reduced by roughly 50 percent in all three strains. Conversely, the strains each produced significantly more pantothenate. This result is also consistent with the model that the two pathways produce competing substrates for PanC (α -HIV and pantoate). Taken together, the above results further indicate that decreasing pantoate synthesis should be beneficial in promoting HMBPA production as well as reducing pantothenate levels.

20

Example IV. Methods for Regulating HMBPA:Pantothenate Levels

As demonstrated in Examples I and II, PanE1 and/or PanE2 contribute to enhanced HMBPA production as does reduced PanB activity. This Example demonstrates that overexpressing PanE1 increases HMBPA production relative to pantothenate production whereas overexpressing PanB decreases HMBPA production relative to pantothenate production. Furthermore, in strains overexpressing IlvC, HMBPA production is enhanced.

PA668 is a derivative of PA824 that contains extra copies of *P*₂₆ *panB* amplified at the *vpr* or *panB* locus. PA668 was constructed using the *panB* expression vector (pAN636) which allows for selection of multiple copies using chloramphenicol. The sequence of pAN636 is set forth as SEQ ID NO:22 and the vector is depicted schematically in Figure 9. The pAN636 *NotI* restriction fragment, missing the *E. coli* vector sequences, was ligated and then used to transform PA824 with selection on plates containing 5 μ g/ml chloramphenicol. Transformants resistant to 30 μ g/ml

30

chloramphenicol were isolated and screened for pantothenate production in 48 hour test tube cultures. The isolates shown produce less HMBPA than PA824 (conversely producing about 10 percent more pantothenate than PA824). A second strain, called PA669, was constructed which is PA824 with extra copies of *P₂₆ panE1* amplified at the *vpr* or *panE1* locus. Strain PA669 was constructed by transforming PA824 with the self-ligated *NotI* fragment of plasmid pAN637 with selection for resistance to chloramphenicol. The sequence of pAN637 is set forth as SEQ ID NO:23 and the vector is depicted schematically in Figure 10. Two isolates of PA669 were chosen for further study; PA669-5 produces less PanE1 than PA669-7 as judged by SDS-PAGE analysis of total cell extracts made from the two strains.

Test tube cultures of strains PA824, PA668-2, PA668-24, and the two isolates of PA669 (PA669-5 and PA669-7) were grown in three different media (SVY, SVY + aspartate, and SVY + aspartate + pantoate) for 48 hours and then assayed for pantothenate, HMBPA, and β -alanine (Table 3).

Table 3. Effect of extra copies of *panB* and *panE1* on pantothenate and HMBPA production by PA824, 48 hour test tube culture data, SV medium.

Strain	<i>panB</i> plasmid	<i>panE</i> plasmid	no additions			+ aspartate (10 g/L)			+ aspartate (10 g/L) & pantoate (5 g/L)		
			[pan] (g/L)	[β-ala] (g/L)	HMBPA *	[pan] (g/L)	[β-ala] (g/L)	HMBPA	[pan] (g/L)	[β-ala] (g/L)	HMBPA
PA824	NONE	NONE	1.8	0.05	<0.1	4.7	2.5	0.53	5.6	2.5	<0.10
PA668-2	pAN636	NONE	1.5	<0.04	<0.1	5.0	1.6	<0.10	4.9	1.2	<0.10
PA668-24	pAN636	NONE	1.8	0.05	<0.1	4.9	2.8	0.34	6.1	2.6	<0.10
PA669-5	NONE	pAN637	1.8	0.04	<0.1	4.2	3.1	0.74	5.8	2.6	0.30
PA669-7	NONE	pAN637	1.8	0.06	<0.1	3.7	3.2	1.41	5.2	2.5	0.75

5 * HMBPA = peak area X 10⁻³

None of the strains produced detectable quantities of HMBPA in SVY medium. All strains produced roughly equivalent amounts of pantothenate and low amounts of β -alanine indicating that β -alanine is limiting for both pantothenate and HMBPA synthesis in these cultures and that β -alanine is a precursor for both compounds. When grown in SVY + aspartate medium, the two PA669 isolates produced more HMBPA than PA824 whereas both PA668 isolates produced less HMBPA than PA824. It is noteworthy that the strain that produces the most PanE1 (PA669-7) produced the most HMBPA (and the least pantothenate). This suggests that high levels of PanE1 favor the production of HMBPA at the expense of lower pantothenate synthesis. It is also interesting that PA668-24 produced more HMBPA than PA668-2, even though SDS-PAGE analysis of extracts from the two strains showed that they produce roughly equivalent levels of PanB. The SDS-PAGE analysis also showed that PA668-24 makes much more IlvC than PA668-2. Based on these data, it is proposed that IlvC influences HMBPA synthesis by increasing steady state levels of α -KIV and/or by catalyzing α -HIV formation from α -KIV, thereby accounting for the observed shift towards production of HMBPA.

The final set of data in Table 3 shows that adding pantoate to the growth medium decreased HMBPA production by all strains that had previously produced detectable levels, *e.g.*, by shifting synthesis towards pantothenate. This further supports the model that α -HIV and pantoate are competitive substrates for PanC.

Example V: Increasing HMBPA Production by Limiting Serine Availability

It was hypothesized that the ratio of pantothenate to HMBPA production could also be controlled by regulating the availability of serine or methylene tetrahydrofolate in the microorganism cultures. In particular, it is proposed that decreasing the availability of serine could increase HMBPA production relative to pantothenate production, whereas increasing the availability of serine would decrease the production of HMBPA relative to pantothenate production. This method is based on the understanding that the PanB substrate, methylenetetrahydrofolate is derived from serine. Thus, regulating serine levels should effectively regulate PanB substrate levels. To test this hypothesis, PA824 was grown in test tube cultures of SVY glucose plus 5 g/L β -alanine and \pm 5 g/L serine for 48 hours at 43°C.

Table 4: *Production of HMBPA and pantothenate by PA824 with and without the addition of serine*

serine added at 5 g/L	OD ₆₀₀	[pan] g/L	[HMBPA] g/L
-	16.3	4.9	0.84
-	14.0	4.5	0.80
+	13.1	6.4	0.56
+	12.9	6.0	0.62

5

As demonstrated in Table 4, addition of serine decreases the level of production while conversely increasing pantothenate production. At least one method of decreasing methylene tetrahydrofolate levels in order to regulate HMBPA production levels is to decrease the activity of serine hydroxymethyl transferase (the *glyA* gene product), thereby decreasing methylene tetrahydrofolate biosynthesis in appropriately engineered microorganisms. At least one method of decreasing serine levels in order to regulate HMBPA production is to decrease the activity of 3-phosphoglycerate dehydrogenase (the *serA* gene product).

15

Example VI: Increasing HMBPA Production by Modifying Culture Conditions for Recombinant Microorganisms

In at least one fermentation (Fermentation P162), levels of HMBPA production reached 35 g/L. Briefly, fermentation of strain PA824 was carried out as described in Example I but utilizing PFM-155 medium formulated as follows.

20

BATCH

	MATERIAL	g/L (final)
1	Amberex 1003	5
2	Cargill 200/20 (soy flour)	40
3	Na Glutamate	5
4	(NH ₄) ₂ SO ₄	8
5	MgSO ₄ ·7H ₂ O	1

6	MAZU DF204C	1
7	H ₂ O	qs to 4 L

Added After Sterilization and Cool Down

1	KH ₂ PO ₄	10
2	K ₂ HPO ₄ ·3H ₂ O	20
3	H ₂ O	qs to 400 ml
1	80% Glucose	20
2	CaCl ₂ ·2H ₂ O	0.1
1	Sodium Citrate	1
2	FeSO ₄ ·7H ₂ O	0.01
3	SM-1000X	1 X

FEED

	MATERIAL	g/L (final)
1	80% Glucose	800
2	CaCl ₂ ·2H ₂ O	0.8
3	H ₂ O	qs to 3500 ml

Added After Sterilization and Cool Down

1	Sodium Citrate	2.0
2	FeSO ₄ ·7H ₂ O	0.02
3	SM-1000X	2 X
4	Glutamate Na	5.0
5	H ₂ O	qs to 500 ml

5

However, as a result of loss of process control during the fermentation, the dissolved oxygen became limiting between 16 and 17 hours and glucose began to accumulate after 16 hours.

These changes in fermentation conditions produced the following significant results at or after 16 hours. Namely, synthesis of HMBPA began to increase with a corresponding decrease in pantothenate synthesis. In the four hour interval before 16 hours the culture produced 7 g/l HMBPA, four hours afterwards, 9.0 g/l.

- 5 Pantothenate was the reverse with 10 g/l and 6.0 g/l produced between 12-16 hours and 16-20 hours, respectively. Between 20 and 36 hours the average rate of HMBPA synthesis was 1.0 g/l hr. Overall, fermentation P162 produced 35 g/l of HMBPA in 36 hours.

- 10 Thus, it appears that overfeeding of glucose, and/or limitation of dissolved oxygen (*e.g.*, beginning at about 16 hours) leads to an increase in HMBPA production. Accordingly, two methods for increasing HMBPA production (relative to pantothenate production) are to increase steady state glucose levels and/or decrease steady state dissolved oxygen levels.

15

- Equivalents Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the
- 20 following claims.

What is claimed:

1. A process for the production of 3-(2-hydroxy-3-methyl-
butyrylamino)-propionic acid (HMBPA), comprising culturing a microorganism under
5 conditions such that HMBPA is produced and detecting the HMBPA produced by said
microorganism.
2. A process for the production of 3-(2-hydroxy-3-methyl-
butyrylamino)-propionic acid (HMBPA), comprising culturing a microorganism under
10 conditions such that HMBPA is produced and isolating the HMBPA produced by said
microorganism.
3. A process for the production of 3-(2-hydroxy-3-methyl-
butyrylamino)-propionic acid (HMBPA), comprising culturing a microorganism having
15 increased keto reductase activity or increased pantothenate synthetase activity in the
presence of excess α -ketoisovalerate and excess β -alanine, such that HMBPA is
produced.
4. A process for the production of 3-(2-hydroxy-3-methyl-
20 butyrylamino)-propionic acid (HMBPA), comprising culturing a microorganism having
increased keto reductase activity and increased pantothenate synthetase activity in the
presence of excess α -ketoisovalerate and excess β -alanine, such that HMBPA is
produced.
5. The process of claim 3 or 4, wherein said microorganism
25 comprises a modified *panE* gene.
6. The process of claim 5, wherein the *panE* gene is overexpressed,
deregulated or present in multiple copies.
7. The process of claim 3 or 4, wherein said microorganism
30 comprises a modified *panE1* gene.
8. The process of claim 3 or 4, wherein said microorganism
comprises a modified *panE2* gene.

9. The process of claim 3 or 4, wherein said microorganism comprises a modified *panE1* gene and a modified *panE2* gene.

5 10. The process of claim 3 or 4, wherein said microorganism comprises a modified *panC* gene.

11. The process of claim 3 or 4, wherein the *panC* gene is overexpressed, deregulated or present in multiple copies.

10

12. The process of claim 3 or 4, wherein said microorganism further has increased acetohydroxyacid isomeroreductase activity.

13. A process for the production of 3-(2-hydroxy-3-methyl-
15 butyrylamino)-propionic acid (HMBPA), comprising culturing a microorganism having increased acetohydroxyacid isomeroreductase activity in the presence of excess α -ketoisovalerate and excess β -alanine, such that HMBPA is produced.

14. The process of claim 12 or 13, wherein said microorganism
20 comprises a modified *ilvC* gene.

15. The process of claim 14, wherein the *ilvC* gene is overexpressed, deregulated or present in multiple copies.

25 16. The process of any one of claims 3, 4 or 11, wherein said microorganism further has reduced ketopantoate hydroxymethyltransferase activity.

17. The process of claim 16, wherein said microorganism comprises a modified *panB* gene.

30

18. The process of claim 16, wherein said microorganism has been deleted for the *panB* gene.

19. A process for the production of 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA), comprising culturing a microorganism having reduced ketopantoate hydroxymethyltransferase activity in the presence of excess α -ketoisovalerate and excess β -alanine, such that HMBPA is produced.

5

20. A method for enhancing production of 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA) relative to pantothenate, comprising culturing a recombinant microorganism under conditions such that the HMBPA production is enhanced relative to pantothenate production.

10

21. A process for the production of 2-hydroxyisovaleric acid (α -HIV), comprising culturing a microorganism which overexpresses PanE1 or PanE2 and which further has reduced PanC or PanD activity under conditions such that α -HIV is produced.

15

22. A process for the production of 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA), comprising culturing a recombinant microorganism having decreased expression or activity of *serA* or *glyA* under conditions such that HMBPA is produced.

20

23. A process for the production of 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA), comprising culturing a recombinant microorganism having decreased expression or activity of *serA* and *glyA* under conditions such that HMBPA is produced.

25

24. The process of any one of the proceeding claims wherein the microorganism is cultured under conditions of increased steady state glucose.

25. The process of any one of the proceeding claims wherein the microorganism is cultured under conditions of decreased steady state dissolved oxygen.

30

26. The process of any one of the proceeding claims wherein the microorganism is cultured under conditions of decreased serine.

27. A product produced according to any one of the above claims.

28. A recombinant microorganism that produces 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA), the microorganism having a modification in at least one gene encoding ketopantoate reductase that results in increased reductase activity and having a mutation or deletion in the *panB* gene that results in reduced ketopantoate hydroxymethyltransferase activity.

29. The recombinant microorganism of claim 28, wherein the gene encoding ketopantoate reductase is a *panE* gene.

10

30. The recombinant microorganism of claim 29, wherein the *panE* gene is *panE1*.

31. The recombinant microorganism of claim 29, wherein the *panE* gene is *panE2*.

15

32. The recombinant microorganism of claim 28, wherein the microorganism has a modification in *panE1* and *panE2*.

33. The recombinant microorganism of claim 28, further having a modification in *ilvC* that results in increased acetohydroxyacid isomeroreductase activity.

20

34. A recombinant microorganism that produces 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA), the microorganism having a modification in *ilvC* that results in increased acetohydroxyacid isomeroreductase activity and having a mutation or deletion in the *panB* gene that results in reduced ketopantoate hydroxymethyltransferase activity.

25

35. The recombinant microorganism of any one of claims 28 to 34, wherein said microorganism belongs to the genus *Bacillus*.

30

36. The recombinant microorganism of claim 35, which is *Bacillus subtilis*.

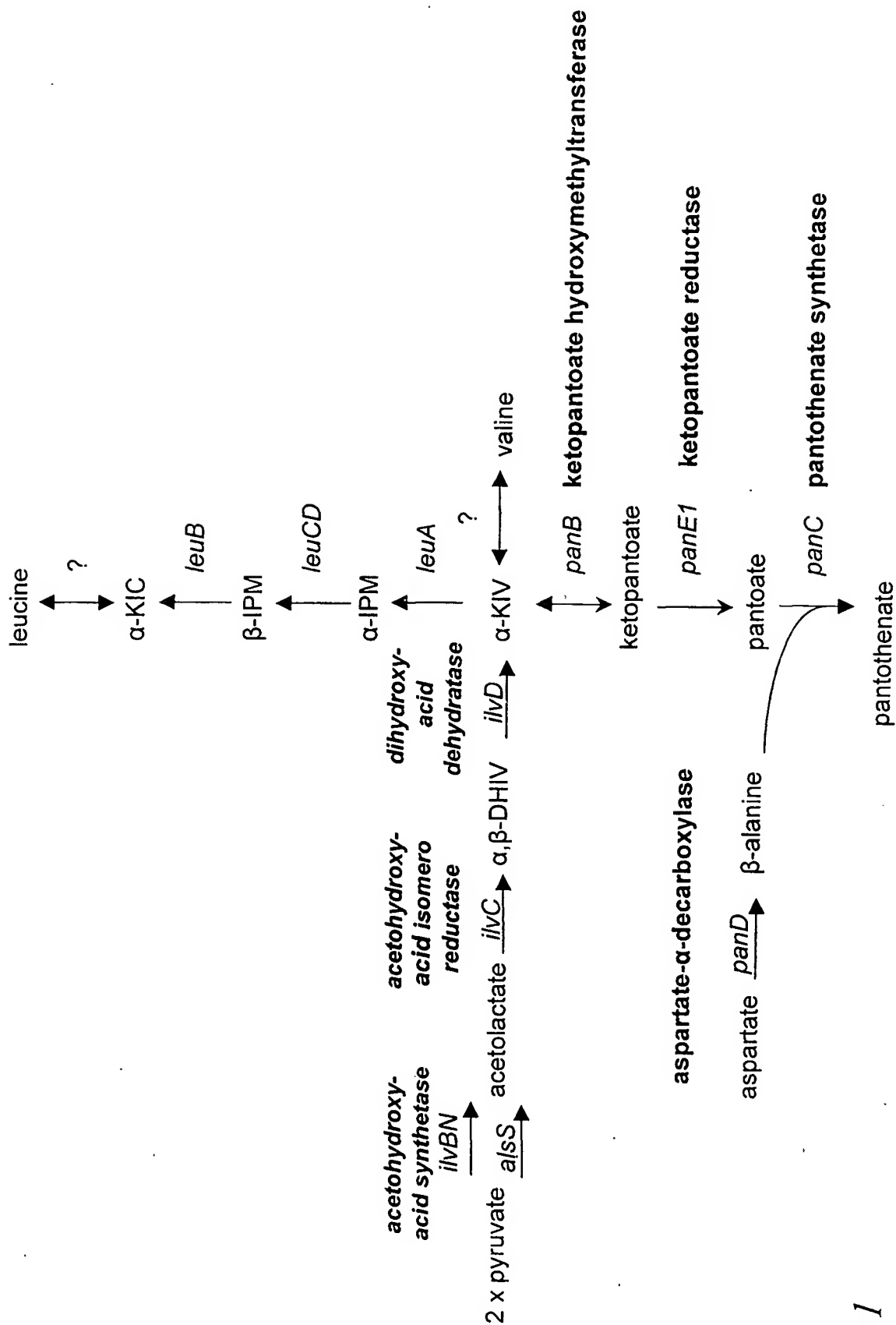
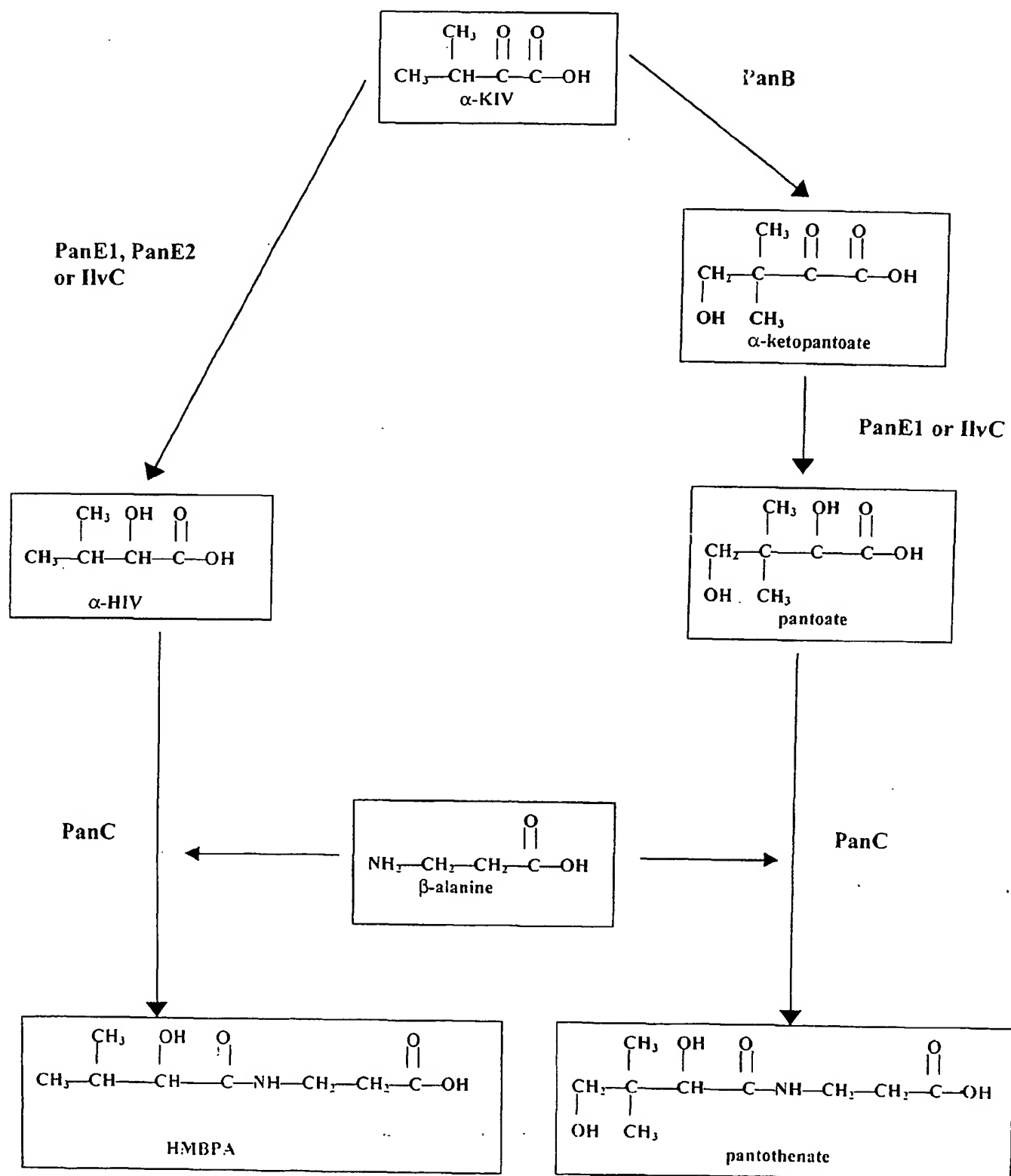
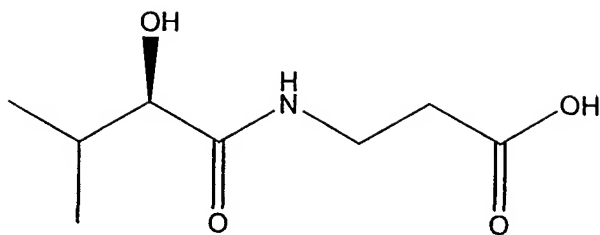


FIG 1

Figure 2. Proposed pathway for biosynthesis of HMBPA.





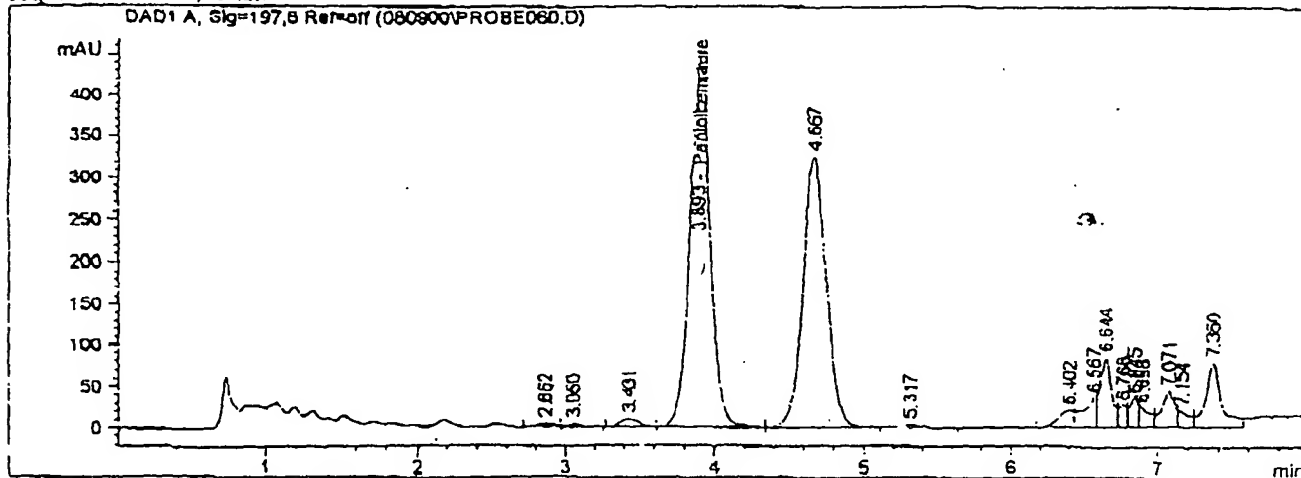
[R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid ("HMBPA")

FIG 3

```

=====
Injection Date   :                               Seq. Line :   42
Sample Name     :   8                           Vial       :   41
Acq. Operator   :                               Inj         :    1
                                           Inj Volume  : 5 ul

Sequence File   : C:\HPCHEM\2\SEQUENCE\VIT-B2.S
Method          : C:\HPCHEM\2\METHODS\PANTOTHE.M
Last changed    :
Aqua 150*2,0mm
  
```



```

=====
External Standard Report
=====
  
```

```

Sorted By           :      Signal
Calib. Data Modified :
Multiplier          :      1.0000
Dilution            :      1.0000
  
```

Signal 1: DAD1 A, Sig=197,6 Ref=off

RetTime [min]	Type	Area [mAU*s]	Amt/Area	Amount [mg/L]	Grp	Name
3.893	VV	4173.96582	9.76719e-2	407.67936		Pantothensäure

Totals : 407.67936

Results obtained with enhanced integrator!

```

=====
*** End of Report ***
=====
  
```

HPLC - analysis of Fermentation broth

Figure 4

SPEC:

Samp: Panthotensre.nebenprodukt Probe 4
 Comm: LCMS Aqua v. Hr. Seb. H2O/ACN 100:1el NH4OAc PH 4.5 197 nm
 Oper: Study:
 Base: 172.10 Masses: 115.00 > 949.99
 Peak: 1000.0 mmu Intensity: 742914
 Scan 168 @ 5.59 min (ESI +QIMS APICID LMR UP PROF) (SM 5)

Scans: 1 > 904

Client: Dr.C.Beck
 #Peaks: 8349
 RIC: 1575177

5.5E+05

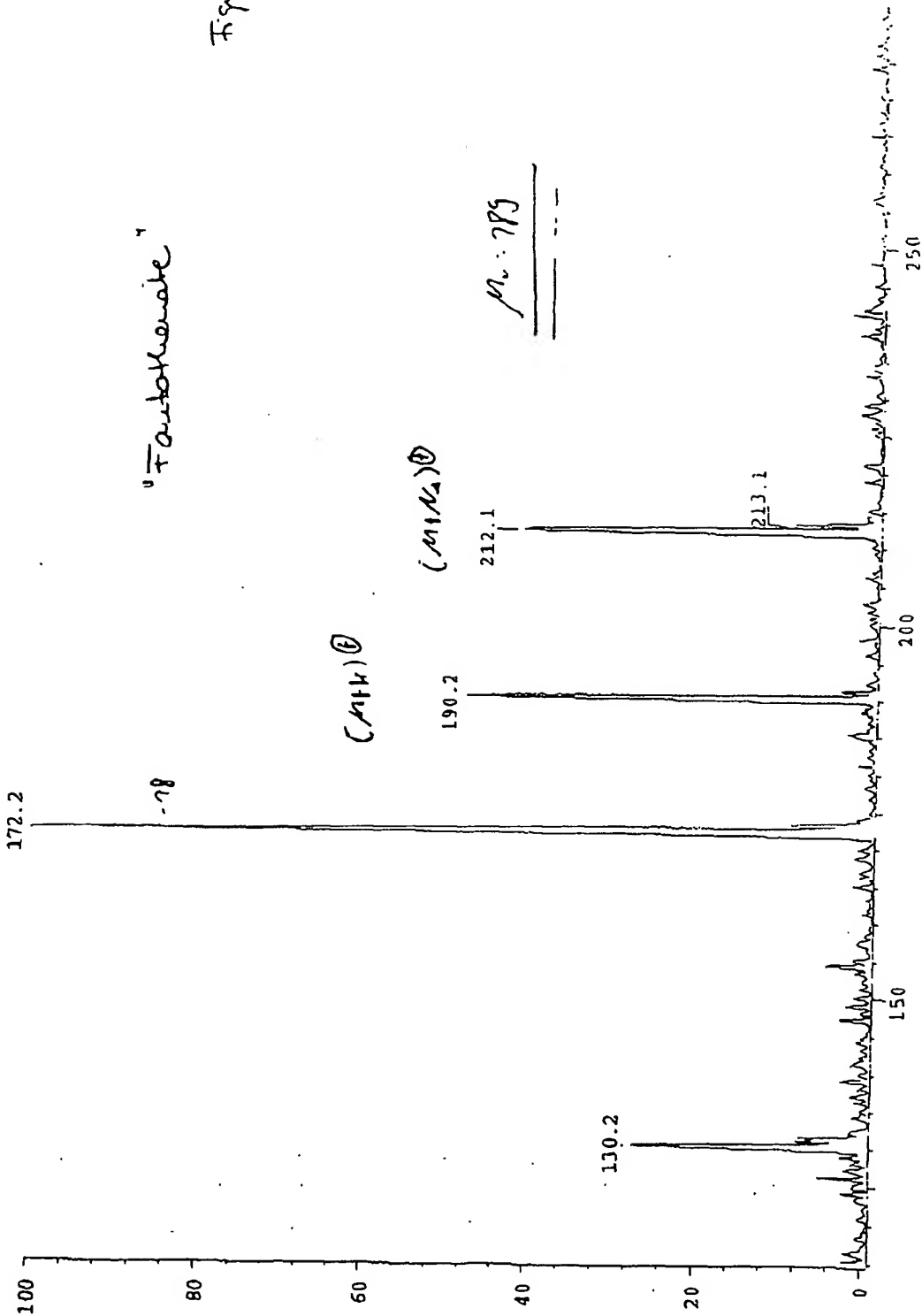


Figure 6. An alignment of the C-terminal amino acids from all known or suspected PanB proteins.

	1 ↓	2 ↓	3 ↓	
CAA10222.1	AYVAE	VGVI	FFGA-EHGFSA	
APE0676	RYAED	VRNGR	FFGE-EHWH-AKEPLEDIS	
RYP04152	LYTE	EV	DGGIYFAE-EHIFQ	
REC04341	QYMA	EVE	SGVYFGE-EHSFH	
sp Q09672	RYTY	EVE	EQGLYFAE-EHSFQ	
RNG01188	AYVA	EV	KAKTFFAA-EHIFAD	
RNM00107	AYVA	EV	KAKTFFAA-EHIFAD	
RPA02174	AYVRA	V	KDVSEFFAA-EHGFNA	
RCA00999	SYAKE	VR	REGTFDE-EHSEK-IDQSIIDETTK	
REF01843	KYTE	EV	KDGVFFGP-EHGFK-ISDQVLEKLY	
RCJ02253	KYRDE	EV	KSGLFFSQ-EHSFDYLDDELLDKLY	
RES02239	GYVQ	D	VRHRAFFEQ-EHSFQ-MNQTVLDGLYGCK	
RDR03436	HYAA	EV	RAREFFSK-DNSEFV-MKDEVLDKLY	
RCY14036	KFSGE	VR	QRFERR-G	
RPA08114	RFAED	VR	RERRFFEA-RHCFAMRE	
RRC02991	AYAA	EV	RSRAFFAP-EFSFDEVKK	
SOC75A.02	AYAED	V	VGGTFFAD-EHSHH	
CAA65397.1	AYTAD	I	HAGTFFGE-AESF	
CAB56202.1	EWAA	A	EKLN-	
RMT01063	QYAQ	EV	AGGVFFAD-EHSF	
sp Q10505	QYAQ	EV	AGGVFFAD-EHSF	
RML00370	QYAE	EV	ASAVFFAE-EHCF	
RML00622	QYAE	EV	ASAVFFAE-EHCF	
FAA01082	NFKI	D	VEGGNFFSE-EESYG	
RHP00462	QYAD	D	VKKGNFFNE-LESYH	
RPG00121	HYTA	D	VKSMDFFNK-DECV	
RPH00184	TFRE	EV	KEGKFFGR-EHWEFQDKEEFKRIKDNVMKKLNL	
RF700767	EFK	EV	KEGKFFGR-EHWEFQDKEEFKRIKDNVMKKLNL	
TML728	EFRE	EV	KEGKFFTE-EHSETDKSKGVSS	
PAB0570	EFK	D	VKEGKFFGR-EHWEFQDKEEFKRIKESVLRKVD	
RSC03568	EYTA	S	VEDRTFFERGTHIS-KVKEDLNNEFLSSINEK	
AAD37248.1	QYRE	EV	KSRAYFAE-QHLYPIPKDELVEFQKAVDELPEEK	
T3F17.24	EYKE	EV	SKVFFGP-SHSPYKITASELDGLFTELQKLGFDKAASAAALAPENMEFSK	

Figure 7. Construction of pAN624.

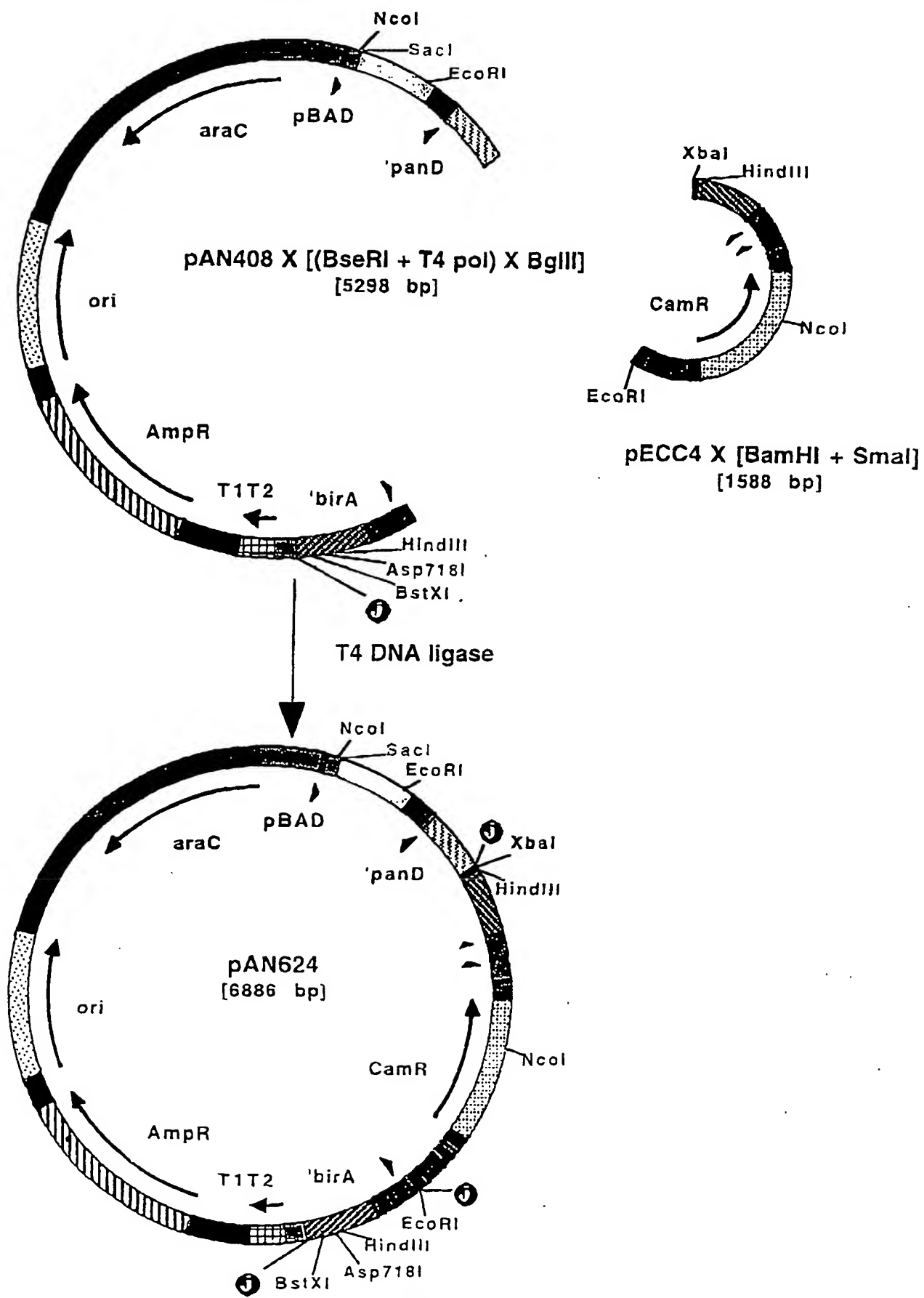


Figure 8. Construction of pAN620.

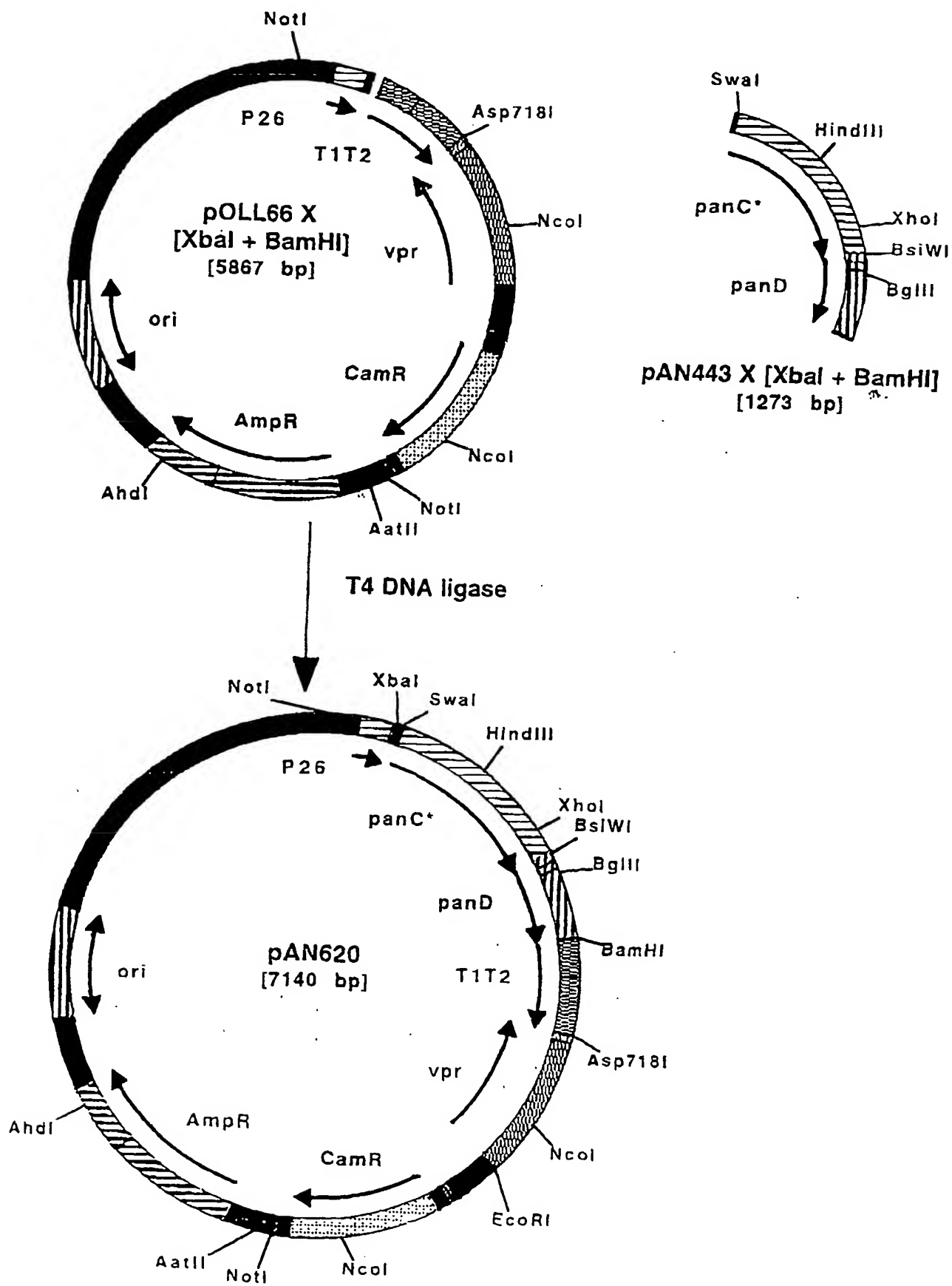


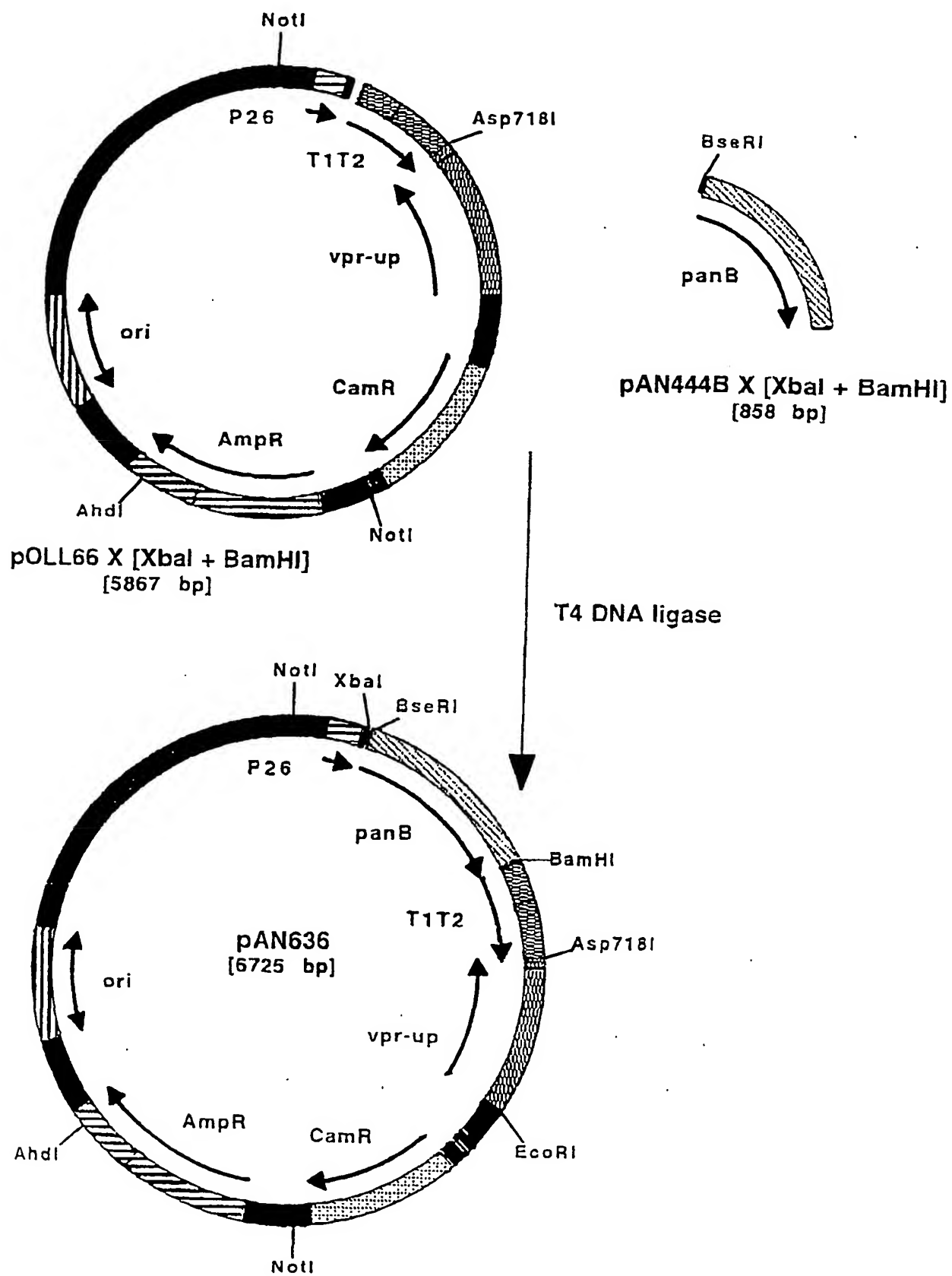
Figure 9 . Construction of pAN636.

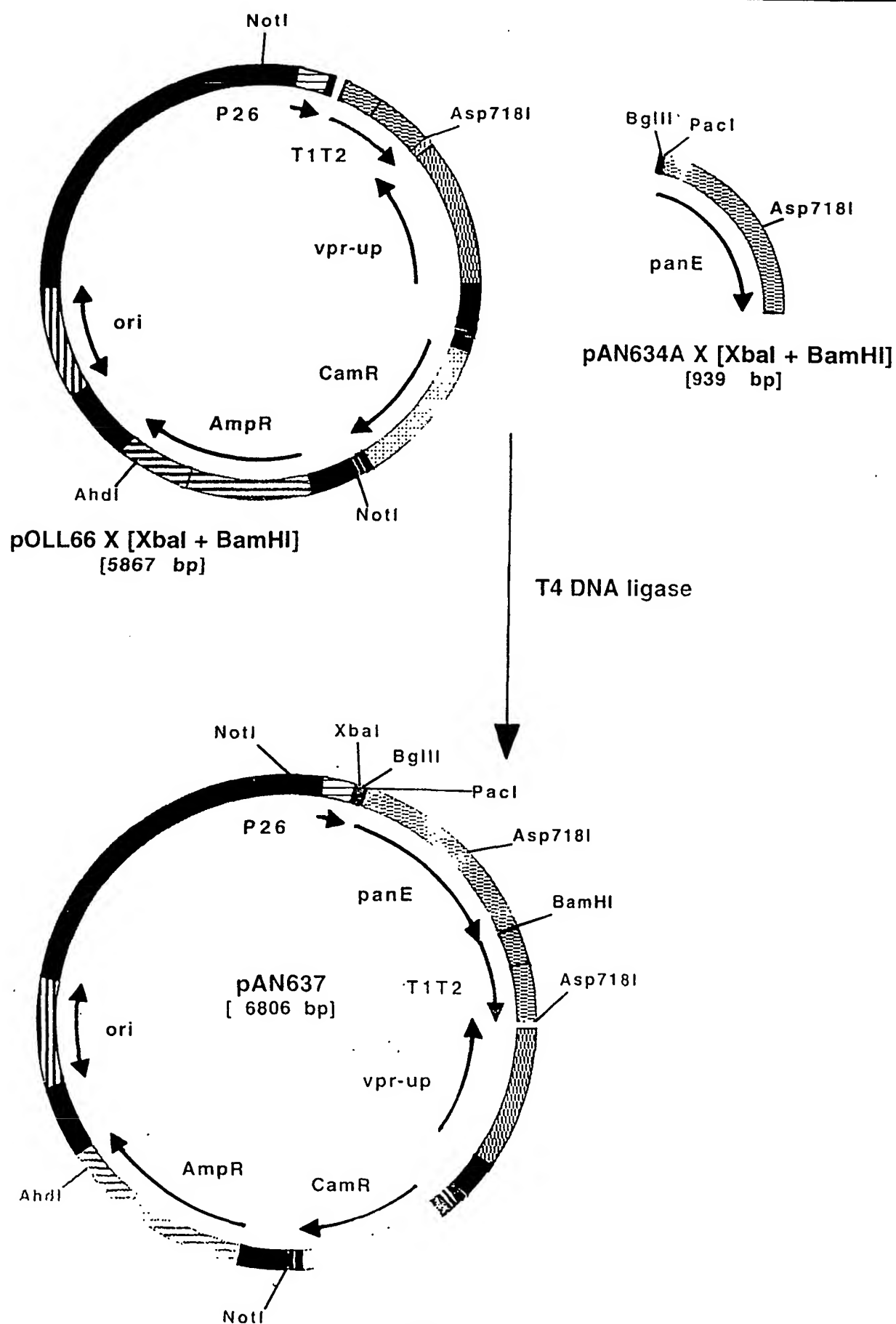
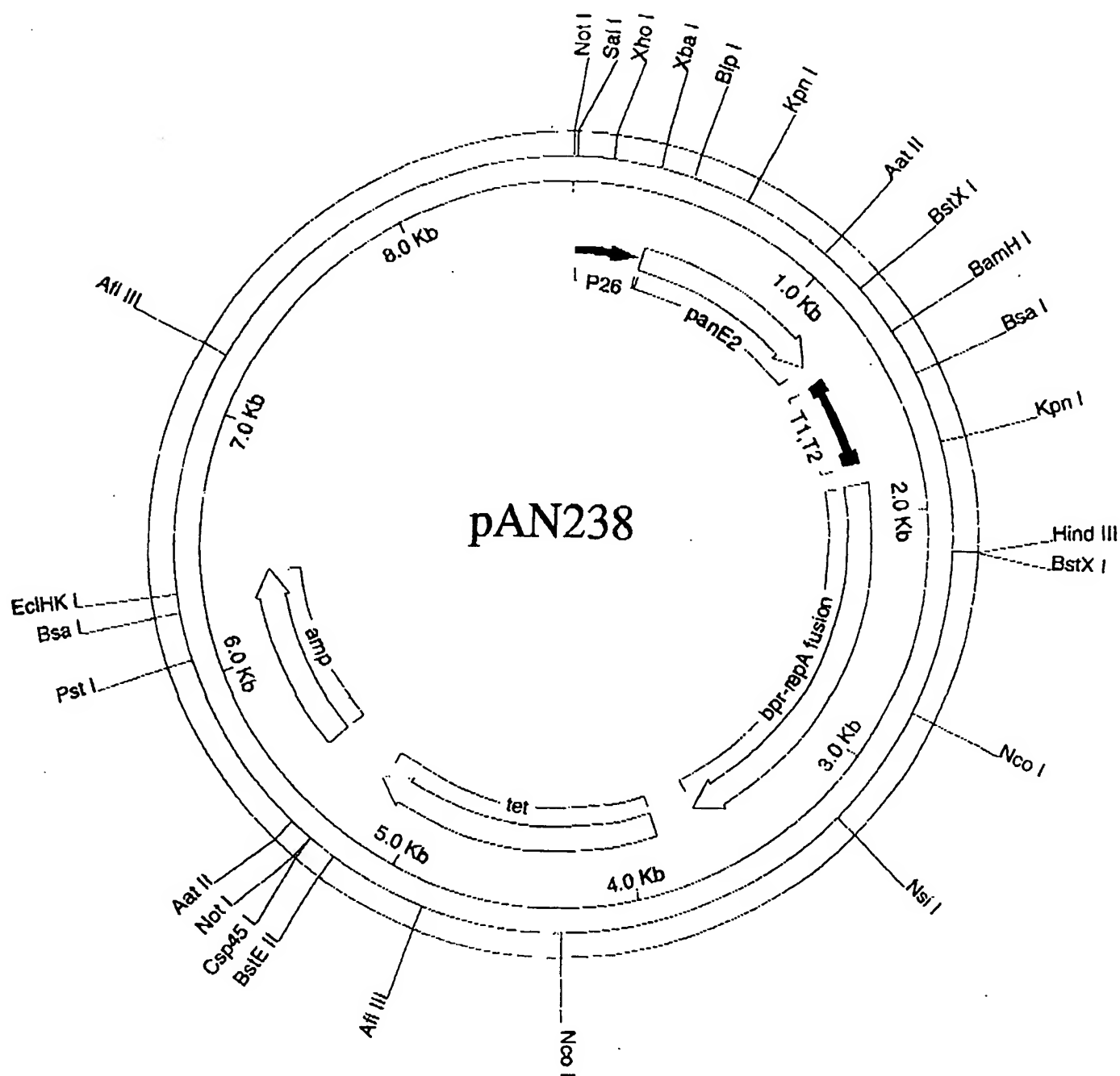
Figure 10 . Construction of pAN637.

Figure 11: Structure of pAN238, a plasmid designed to overexpress the panE2 gene from the P₂₆ promoter after integrating at bpr or panE2.



SEQUENCE LISTING

<110> OmniGene BioProducts Inc., *et al.*

<120> METHODS AND MICROORGANISMS FOR THE PRODUCTION OF 3-(2-HYDROXY-3-METHYL-BUTYRLAMINO)-PROPIONIC ACID (HMBPA)

<130> BGI-146PC

<140>

<141>

<160> 25

<170> PatentIn Ver. 2.0

<210> 1

<211> 194

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:promoter sequence

<220>

<221> -35_signal

<222> (136)..(141)

<220>

<221> -10_signal

<222> (159)..(164)

<400> 1

gctattgacg acagctatgg ttcactgtcc accaaccaaa actgtgctca gtaccgcaa 60

tatttctccc ttgaggggta caaagagggtg tccctagaag agatccacgc tgtgtaaaaa 120

ttttacaaaa aggtattgac tttccctaca ggggtgtgtaa taatttaatt acaggcgggg 180

gcaaccccgct ctgt

194

<210> 2

<211> 163

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:promoter sequence

<220>

<221> -35_signal

<222> (113)..(118)

<220>

<221> -10_signal

<222> (136)..(141)

<400> 2
gcctacctag cttccaagaa agatattccta acagcacaag agcggaaaga tgttttgttc 60
tacatccaga acaacctctg ctaaaattcc tgaaaaattt tgcaaaaagt tgttgacttt 120
atctacaagg tgtggtataa taatcttaac aacagcagga cgc 163

<210> 3
<211> 127
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:promoter
sequence

<220>
<221> -35_signal
<222> (34)..(39)

<220>
<221> -10_signal
<222> (58)..(63)

<220>
<221> -35_signal
<222> (75)..(80)

<220>
<221> -10_signal
<222> (98)..(103)

<400> 3
gaggaatcat agaattttgt caaaataatt ttattgacaa cgtcttatta acgttgatat 60
aatTTaaatt ttatttgaca aaaatgggct cgtgttgtag aataaatgta gtgagggtgga 120
tgcaatg 127

<210> 4
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:ribosome
binding site

<400> 4
taaacaatgag gaggagaaaa catg 24

<210> 5
<211> 28
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome
binding site

<400> 5

attcgagaaa tggagagaat ataatatg

28

<210> 6

<211> 13

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome
binding site

<400> 6

agaaaggagg tga

13

<210> 7

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome
binding site

<220>

<223> All occurrences of n = any nucleotide

<400> 7

ttaagaaagg aggtgannnn atg

23

<210> 8

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome
binding site

<220>

<223> All occurrences of n = any nucleotide

<400> 8

ttagaaagga ggtgannnnn atg

23

<210> 9

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome
binding site

<220>

<223> All occurrences of n = any nucleotide

<400> 9

agaaaggagg tgannnnnnn atg

23

<210> 10

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome
binding site

<220>

<223> All occurrences of n = any nucleotide

<400> 10

agaaaggagg tgannnnnna tg

22

<210> 11

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome
binding site

<400> 11

ccctctagaa ggaggagaaa acatg

25

<210> 12

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome
binding site

<400> 12

ccctctagag gaggagaaaa catg

24

<210> 13

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome
binding site

<400> 13

ttagaaagga ggatttaa atg

23

<210> 14

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome binding site

<400> 14

ttagaaagga ggtttaatta atg

23

<210> 15

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome binding site

<400> 15

ttagaaagga ggtgatttaa atg

23

<210> 16

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome binding site

<400> 16

ttagaaagga ggtgttttaa atg

23

<210> 17

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome binding site

<400> 17

attcgagaaa ggaggtgaat ataatatg

28

<210> 18

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome binding site

<400> 18

attcgagaaa ggaggtgaat aataatg

27

<210> 19

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: ribosome binding site

<400> 19

attcgtagaa aggaggtgaa ttaatatg

28

<210> 20

<211> 6886

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: vector

<220>

<223> pAN624

<400> 20

```

aagaaaccaa ttgtccatat tgcacagac attgccgtca ctgcgtcttt tactggctct 60
tctcgtaac caaaccggtta acccgctta ttaaaagcat tctgtaacaa agcgggacca 120
aagccatgac aaaaacgcgt aacaaaagt tctataatca cggcagaaaa gtccacattg 180
attatttgca cggcgtcaca ctttgctatg ccatagcatt tttatccata agattagcgg 240
atcctacctg acgcttttta tgcgaactct ctactgtttc tccataaccg ttttttggg 300
ctaacaggag gaattaacca tggatccgag ctgcacagta tcaagcactt cacaatctgg 360
gagctgaaag cccgccttat gtagctcata cttgacaaat ccaagggtcaa aatggatatt 420
gtgggcgaca aaataagcgc cgtcaagcaa ttggaatact tcttcagcaa ctgcttcaaa 480
tggtgttca ttctcgacca tttgattaga gattccagta agctgctcaa taaaagcagg 540
gattgattta tttggattaa tgtattttga aaaccgctca gtaatttgtc cattttcgat 600
tacaaccgct gcgatttgta tgattttatc gcctttcttc ggcgaattcc ctgttgtctc 660
tacatctata acaacgaacc gttgcttatt cattaatatg gacacctcaa ttcttgcata 720
cgacaaaagt gtaacacgtt ttgtacggaa atggagcggc aaaaccggtt tactctcaaa 780
atcttaaaag aaaacccccg ataaagggg cttttcttct acaaaattgt acgggctgg 840
tcgttccccg gcatttgctt aattttgttt tgatcattca gaacagccac ttctcgctca 900
tggttgccg cttcttgatc agacatcatt ttgtaggaaa taataatgac cttatctcct 960
tctgcacaa ggctgctggc tgcaccgttt aagcatatga cgcgcttcc ccgtttacca 1020
ggaataatat acgtttcaag acgtgctcca ttattattat tcacaatttg tactttttca 1080
ttaggaagca ttcccacagc atcaatgaga tctctagag tgcacctgca ggcagcaag 1140
cttccgtcga cgtctcctt tatgcgactc ctgcattagg aagcagccca gtagtaggtt 1200
gaggccgttg agcaccgctg ccgcaaggaa tgggtcatgc aaggagatgg cgccaacag 1260
tcccccgcc ccggggcctg ccaccatacc cagccgaaa caagcgtca tgagcccgaa 1320
gtggcgagcc cgtcttccc catcggtgat gtgcggcata taggcgccag caaccgcacc 1380
tgtggcgccg gtgatgccg ccacgatgcg tccggcgtag aggatcaatc ttcatccatt 1440
ccaaggtaaa tcccccttgc ccgtttctgt taccattata caccttttga accttaacgt 1500
aaacgttaag ttttaaaaaa caataaaaaa gacgagcagc atacagcacc cgtctttcac 1560
tttctgttt aagctaaact tcccgcact gacagagact ctttttgaag gctttcagaa 1620
agcactcgat acgcgatctg gagctgtaat ataaaaacct tcttcaacta acggggcagg 1680
ttagtgacat tagaaaaccg actgtaaaaa gtacagtcgg cattatctca tattataaaa 1740
gccagtcatt aggcctatct gacaattcct gaatagagtt cataaacaat cctgcatgat 1800
aaccatcaca aacagaatga tgtacctgta aagatagcgg taaatatatt gaattacctt 1860
tattaatgaa ttttctgtct gtaataatgg gtagaaggta attactatta ttattgatatt 1920
ttaagttaaa ccagtaaat gaagtccatg gaataataga aagagaaaaa gcattttcag 1980
gtataggtgt tttgggaaac aatttccccg aaccattata tttctctaca tcagaaagg 2040
ataaatcata aaactctttg aagtcattct ttacaggagt ccaaatacca gagaatgttt 2100
tagatacacc atcaaaaatt gtataaagt gctctaactt atcccaataa cctaactctc 2160
cgtcgctatt gtaaccagtt ctaaaagctg tatttgagtt tatcacctt gtcactaaga 2220
aaataaatgc agggtaaaat ttatatcctt cttgttttat gtttcggtat aaaacactaa 2280

```


tatcaatttc	tgtggttata	ctaaaagtcg	tttgttggtt	caaataatga	ttaaatatct	2340
cttttctctt	ccaattgtct	aaatcaatct	tattaaagtt	catttgatat	gcctcctaaa	2400
tttttatcta	aagtgaattt	aggaggctta	cttgtctgct	ttcttcatta	gaatcaatcc	2460
ttttttaaaa	gtcaatatta	ctgtaacata	aatatatatt	ttaaaaatat	cccactttat	2520
ccaattttcg	tttgttgaac	taatgggtgc	tttagttgaa	gaataaagac	cacattaaaa	2580
aatgtggtct	tttgtgtttt	tttaaaggat	ttgagcgtag	cgaaaaatcc	ttttctttct	2640
tatcttgata	ataagggtaa	ctattgcatg	ataagctgtc	aaacatgaga	attcccgttt	2700
tcttctgcaa	gccaaaaaac	cttcggttac	aacgagaagg	attcttcact	ttctaaagtt	2760
cggcgagttt	catccctctg	tcccagtcct	tttttggatc	aaggcagact	gctgcaatgt	2820
ctatctatct	taataatagg	tgcagttcgc	aggcgatact	gcccaatgga	agtataccaa	2880
aatcaacggg	cttgtacca	cacattagcc	caattcgata	tcggcagaat	agattttttt	2940
aatgccttcg	ttcgtttcta	aaagcagaac	gccttcatca	tctataccta	acgccttacc	3000
gtaaaagggt	ccgtttaacg	ttctggctct	catattagtg	ccaataccga	gcgcatagct	3060
ttcccataaa	agcttaatcg	gcgtaaatcc	gtgcgtcata	taatcccggg	accggtttctc	3120
aaagcatagt	aaaatatgct	ggatgacgcc	ggcccgatca	atTTTTTccc	cagcagcttg	3180
gctgaggctt	gtcgcgatgt	ccttcaattc	atctggaaaa	tcattaggct	gctgggttaa	3240
cggctctccag	cttggctggt	ttggcggatg	agagaagatt	ttcagcctga	tacagattaa	3300
atcagaacgc	agaagcggtc	tgataaaaca	gaatttgcct	ggcggcagta	gcgcgggtgg	3360
cccacctgac	cccatgccga	actcagaagt	gaaacgccgt	agcgcgatg	gtagtgtggg	3420
gtctcccat	gcgagagtag	ggaactgcca	ggcatcaaat	aaaacgaaag	gctcagtcga	3480
aagactgggc	ctttcgtttt	atctgttgtt	tgtcggtgaa	cgctctcctg	agtaggacaa	3540
atccgcggg	agcggatttg	aacgttgcga	agcaacggcc	cggagggtgg	cgggcaggac	3600
gcccgcata	aactgccagg	catcaaatta	agcagaaggc	catcctgacg	gatggccttt	3660
ttcggtttct	acaaactctt	tttgtttatt	ttctaaata	cattcaaata	tgtatccgct	3720
catgagacaa	taaccctgat	aaatgcttca	ataatattga	aaaaggaaga	gtatgagtat	3780
tcaacatttc	cgtgtcgccc	ttattccctt	ttttgcggca	ttttgccttc	ctgtttttgc	3840
tcaccagaa	acgctggtga	aagtaaaaga	tgctgaagat	cagttgggtg	cacgagtggg	3900
ttacatcgaa	ctggatctca	acagcggtaa	gatccttgag	agttttcgcc	ccgaagaacg	3960
ttttccaatg	atgagcaact	ttaaagttct	gctatgtggc	gcggtattat	cccgtgttga	4020
cgccgggcaa	gagcaactcg	gtcgcgcgat	acactattct	cagaatgact	tggttgagta	4080
ctcaccagtc	acagaaaagc	atcttacgga	tgccatgaca	gtaagagaat	tatgcagtgc	4140
tgccataacc	atgagtgata	acactgcggc	caacttactt	ctgacaacga	tcggaggacc	4200
gaaggagcta	accgcttttt	tgcacaacat	gggggatcat	gtaactcgcc	ttgatcgttg	4260
ggaaccggag	ctgaatgaag	ccataccaaa	cgacgagcgt	gacaccacga	tgctgttagc	4320
aatggcaaca	acgttgcgca	aactattaac	tggcgaacta	cttactctag	cttcccggca	4380
acaattaata	gactggatgg	aggcggataa	agttgcagga	ccacttctgc	gctcggccct	4440
tccggctggc	tggtttattg	ctgataaatc	tgagcgggt	gagcgtgggt	ctcgcggtat	4500
cattgcagca	ctggggccag	atggtaagcc	ctcccgtatc	gtagtatatc	acacgacggg	4560
gagtcaggca	actatggatg	aacgaaatag	acagatcgct	gagataggtg	cctcactgat	4620
taagcatttg	taactgtcag	accaagttta	ctcatatata	ctttagattg	atttaaaact	4680
tcatttttaa	tttaaaagga	tctaggtgaa	gatccttttt	gataatctca	tgacccaaa	4740
cccttaacgt	gagttttcgt	tccactgagc	gtcagacccc	gtagaaaaga	tcaaaggatc	4800
ttcttgagat	ccttttttct	tgcgcgtaac	ctgctgcttg	caaacaaaaa	aaccaccgct	4860
accagcgtg	gtttgtttgc	cggatcaaga	gctaccaact	ctttttccga	aggtaactgg	4920
cttcagcaga	gcgcagatac	caaatactgt	ccttctagtg	tagccgtagt	taggccacca	4980
cttcaagaac	tctgtagcac	cgcctacata	cctcgctctg	ctaatacctgt	taccagtggc	5040
tgctgccagt	ggcgataagt	cgtgtcttac	cgggttggac	tcaagacgat	agttaccgga	5100
taaggcgcag	cggtcgggct	gaacgggggg	ttcgtgcaca	cagcccagct	tgagcgaac	5160
gacctacacc	gaactgagat	acctacagcg	tgagctatga	gaaagcgcca	cgcttcccga	5220
agggagaaag	gcggacaggt	atccggttaag	cggcagggtc	ggaacaggag	agcgcacgag	5280
ggagcttcca	gggggaaacg	cctggatatct	ttatagtcct	gtcgggtttc	gccacctctg	5340
acttgagcgt	cgatttttgt	gatgctcgtc	agggggcgcg	agcctatgga	aaaacgccag	5400
caacgcggcc	tttttacggt	tccctggcctt	ttgctggcct	tttgctcaca	tgttctttcc	5460
tgcgttatcc	cctgattctg	tggataaccg	tattaccgcc	tttgagttag	ctgataaccgc	5520
tcgcccgcagc	cgaacgaccg	agcgcagcga	gtcagttagc	gaggaaagcg	aagagcgcct	5580
gatgcggtat	tttctcctta	cgcatctgtg	cggtatattca	caccgcataat	ggtgcactct	5640
cagtacaatc	tgctctgatg	ccgcatagtt	aagccagtat	acactccgct	atcgctacgt	5700
gactgggtca	tggtgcgcc	ccgacacccg	ccaacacccg	ctgacgcgcc	ctgacgggct	5760
tgtctgctcc	cggcatccgc	ttacagacaa	gctgtgaccg	tctccgggag	ctgcatgtgt	5820
cagaggtttt	caccgtcatc	accgaaacgc	gcgaggcagc	agatcaattc	gcgcgcgaag	5880

```

gcgaagcggc atgcataatg tgcctgtcaa atggacgaag cagggattct gcaaacccta 5940
tgctactccg tcaagccgtc aattgtctga ttcgttacca attatgacaa cttgacggct 6000
acatcattca ctttttcttc acaaccggca cggaactcgc tcgggctggc cccggtgcat 6060
tttttaataa cccgcgagaa atagagttag tgcgcaaaac caacattgag accgacgggtg 6120
gcgataggca tccgggtggt gctcaaaagc agcttcgcct ggctgatacg ttggtcctcg 6180
cgccagctta agacgcta at ccctaactgc tggcggaata gatgtgacag acgcgacggc 6240
gacaagcaaa catgctgtgc gacgctggcg atatcaaaat tgctgtctgc caggtgatcg 6300
ctgatgtact gacaagcctc gcgtaccgga ttatccatcg gtggatggag cgactcgta 6360
atcgcttcca tgcgccgag taacaattgc tcaagcagat ttatcgccag cagctccgaa 6420
tagcgccctt ccccttgccc ggcgttaatg atttgcccac acaggtcgct gaaatgcggc 6480
tggtgcgctt catccgggag aaagaacccc gtattggcaa atattgacgg ccagttaagc 6540
cattcatgcc agtaggcgag cggacgaaag taaaccact ggtgatacca ttcgcgagcc 6600
tccgatgac gaccgtagtg atgaatctct cctggcgagg acagcaaaat atcaccgggt 6660
cggcaaaaca attctcgtcc ctgatttttc accaccccct gaccgcgaat ggtgagattg 6720
agaatataac ctttcattcc cagcggctcg tcgataaaaa aatcgagata accgttggcc 6780
tcaatcggcg ttaaaccgag caccagatgg gcattaaacg agtatcccg cagcagggga 6840
tcattttgag cttcagccat acttttcata ctcccggcat tcagag 6886

```

<210> 21

<211> 7140

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: vector

<220>

<223> pAN620

<400> 21

```

tcggcggccg cttcgtcgac cgaaacagca gttataaggc atgaagctgt ccggtttttg 60
caaaagtggc tgtgactgta aaaagaaatc gaaaaagacc gttttgtgtg aaaacggctc 120
ttttgtttcc ttttaaccaa ctgccataac tcgaggccta cctagcttcc aagaaagata 180
tcctaacagc acaagagcgg aaagatgttt tgttctacat ccagaacaac ctctgctaaa 240
attcctgaaa aattttgcaa aaagtgtgtg actttatcta caaggtgtgg tataataatc 300
ttaacaacag caggacgctc tagattagaa aggaggattt aaatatgaga cagattactg 360
atatttcaca gctgaaagaa gccataaaac aataccattc agagggcaag tcaatcggat 420
ttgttccgac gatgggggtt ctgcatgagg ggcatttaac cttagcagac aaagcaagac 480
aagaaaacga cgccgttatt atgagtattt ttgtgaatcc tgcacaattc ggccctaattg 540
aagattttga agcatatccg cgcgatattg agcgggatgc agctcttgca gaaaacgccg 600
gagtcgatat tctttttacg ccagatgctc atgatatgta tcccgggtgaa aagaatgtca 660
cgattcatgt agaaagacgc acagacgtgt tatgcgggag ctcaagagaa ggacattttg 720
acggggtcgc gatcgtactg acgaagcttt tcaatctagt caagccgact cgtgcctatt 780
tcgggtttaa agatgcgcag caggtagctg ttgttgatgg gttaatcagc gacttcttca 840
tggatattga attggttcct gtcgatacgg tcagtagagg agacggctta gccaaaagct 900
ctcgcaatgt atacttaaca gctgaggaaa gaaaagaagc gcctaagctg tatcgggccc 960
ttcaaacaag tgcggaactt gtccaagccg gtgaaagaga tcctgaagcg gtgataaaag 1020
ctgcaaaaga tatcattgaa acgactagcg gaaccataga ctatgtagag ctttattcct 1080
atccggaact cgagcctgtg aatgaaattg ctggaaagat gattctcgct gttgcagttg 1140
ctttttcaaa agcgcgttta atagataata tcattattga tattcgtaga aaggaggtga 1200
attaatatgt atcgtacgat gatgagcggc aaacttcaca gggcaactgt tacggaagca 1260
aacctgaact atgtgggaag cattacaatt gatgaagatc tcattgatgc tgtgggaatg 1320
cttcctaattg aaaaagtaca aattgtgaat aataataatg gagcacgtct tgaaacgtat 1380
attattcctg gtaaaccggg aagcggcgtc atatgcttaa acggtgcagc cgcacgcctt 1440
gtgcaggaag gagataaggt cattattatt tcctacaaaa tgatgtctga tcaagaagcg 1500
gcaagccatg agccgaaagt ggctgttctg aatgatcaaa acaaaattga acaaattgctg 1560
gggaacgaac cagcccgtac aattttgtta aggatcctgt tttggcggat gagagaagat 1620
tttcagcctg atacagatta aatcagaacg cagaagcggc ctgataaaac agaatttgcc 1680
tggcggcagt agcgcgggtg tcccacctga ccccatgccg aactcagaag tgaaacgccg 1740
tagcggcagat ggtagtgtgg ggtctcccca tgcgagagta gggaactgcc aggcattcaa 1800

```

taaaacgaaa	ggctcagtcg	aaagactggg	cctttcgttt	tatctgttgt	ttgtcgggtga	1860
acgctctcct	gagtaggaca	aatccgccgg	gagcggattt	gaacggttgcg	aagcaacggc	1920
ccggagggtg	gcgggcagga	cgcccgccat	aaactgccag	gcatcaaatt	aagcagaagg	1980
ccatcctgac	ggatggcctt	tttgcgtttc	tacaaactct	tggtagcgag	acgatcgctcc	2040
tctttgttgt	agcccatcac	ttttgctgaa	gagtaggagc	cgaaagtgcg	ggcgtattca	2100
ttgagcggca	gctgagtcgc	accgacagaa	atcgcttctc	ttgatgtgcc	cggcgatccg	2160
actgtccagc	cgttcgggtc	gctgttgccg	tttgaggtaa	cagcgacaac	gccttctgac	2220
atggcccgat	caagcgctgt	gcttgtcgcc	cagtcggggt	tgtttaaaga	gtttccgaga	2280
gacaggttca	tcacatctgc	cccgctcctgc	actgcacgtt	ccacgcccgc	gatgacgttt	2340
tccgttgtgc	cgcttccgcc	aggccctaac	acacgataag	caagaagtgt	ggcatcaggc	2400
gctacgcctt	taatcgttcc	gtttgcagcc	acagttccgg	ctacgtgtgt	gccatggtca	2460
gttgccctgc	ccctcggatc	gccggttggt	gtttcttttg	gatcgtaatc	attgtccaca	2520
aaatcgatc	ctttatattg	tccaaagttt	ttcttcagat	ctgggtgatt	gtattcaacc	2580
ccagtgtcaa	taatcgccac	cttgatgcct	tttctgtgt	agcctaaatc	ccatgcatcg	2640
tttgctccga	tataaggcgc	actgtcatcc	atgtgaggag	atagggcgtc	ttcggagatt	2700
gtggggaatt	ctcatgtttg	acagcttatc	atgcaatagt	tacccttatt	atcaagataa	2760
gaaagaaaag	gatttttctg	tacgctcaaa	tcctttaaaa	aaacacaaaa	gaccacattt	2820
tttaattgtg	tctttattct	tcaactaaag	caccatttag	ttcaacaaac	gaaaattgga	2880
taaagtggga	tattttttaa	atatatatatt	atgttacagt	aatattgact	tttaaaaaag	2940
gattgattct	aatgaagaaa	gcagacaagt	aagcctccta	aattcacttt	agataaaaaat	3000
ttaggaggca	tatcaaataa	actttaataa	aattgattta	gacaattgga	agagaaaaaga	3060
gatatttaaa	cattatttga	accaacaaac	gacttttagt	ataaccacag	aaattgatata	3120
tagtggttta	taccgaaaca	taaaacaaga	aggatataaa	ttttaccctg	catttatttt	3180
cttagtgaca	agggtgataa	actcaaatac	agcttttaga	actggttaca	atagcgacgg	3240
agagttaggt	tattgggata	agttagagcc	actttataca	atttttgatg	gtgtatctaa	3300
aacattctct	ggtatttgga	ctcctgtaaa	gaatgacttc	aaagagtttt	atgatttata	3360
cctttctgat	gtagagaaat	ataatggttc	ggggaaattg	tttcccaaaa	cacctatacc	3420
tgaaaatgct	ttttctcttt	ctattattcc	atggacttca	tttactgggt	tttaacttaa	3480
tatcaataat	aatagtaatt	accttctacc	cattattaca	gcaggaaaat	tcattaataa	3540
aggtaattca	atatattttac	cgctatcttt	acaggtagat	cattctgttt	gtgatgggtta	3600
tcattgcagga	ttgtttatga	actctattca	ggaattgtca	gataggccta	atgactggct	3660
tttataatat	gagataatgc	cgactgtact	ttttacagtc	gggttttctaa	tgtcactaac	3720
ctgccccggt	agttgaagaa	cgaagcggcc	gcaattcttg	aagacgaaag	ggcctcgtga	3780
tacgcctatt	tttatagggt	aatgtcatga	taataatggt	ttcttagacg	tcagggtggca	3840
cctttcgggg	aaatgtgcgc	ggaaccccta	tttggtttatt	tttctaaata	cattcaaata	3900
tgtatccgct	catgagacaa	taaccctgat	aaatgcttca	ataatattga	aaaaggaaga	3960
gtatgagtat	tcaacatttc	cgtgtcgccc	ttattccctt	ttttgcgga	ttttgccttc	4020
ctgtttttgc	tcacccagaa	acgctggtga	aagtaaaaga	tgctgaagat	cagttgggtg	4080
cacgagtggg	ttacatcgaa	ctggatctca	acagcggtaa	gacccctgag	agttttcgcc	4140
ccgaagaacg	ttttccaatg	atgagcactt	ttaaagttct	gctatgtggc	gcggtattat	4200
cccgtattga	cgccgggcaa	gagcaactcg	gtcgccgcac	acactattct	cagaatgact	4260
tggttgagta	ctcaccagtc	acagaaaagc	atcttacgga	tgccatgaca	gtaagagaat	4320
tatgcagtgc	tgccataaac	atgagtata	acactcgccg	caacttactt	ctgacaacga	4380
tcggaggacc	gaaggagcta	accgcttttt	tgcaacaacat	gggggatcat	gtaactcgcc	4440
ttgatcggtg	ggaaccggag	ctgaatgaag	ccataccaaa	cgacgagcgt	gacaccacga	4500
tgccctgcgc	aatggcaaca	acgttgcgca	aactattaac	tgccgaacta	cttactctag	4560
cttcccgcca	acaattaata	gactggatgg	aggcggataa	agttgcagga	ccacttctgc	4620
gctcgccctt	tccggctggc	tggtttattg	ctgataaatc	tggaagccgg	gagcgtgggt	4680
ctcgcggtat	cattgcagca	ctggggccag	atggtaagcc	ctcccgatc	gtagttatct	4740
acacgacggg	gagtcaggca	actatggatg	aacgaaatag	acagatcgct	gagatagggt	4800
cctcactgat	taagcattgg	taactgtcag	accaagttta	ctcatatata	cttttagattg	4860
attttaaact	tcatttttaa	ttttaaagga	tctaggtgaa	gatccttttt	gataatctca	4920
tgacaaaaat	cccttaacgt	gagttttcgt	tccactgagc	gtcagacccc	gtagaaaaga	4980
tcaaaggatc	ttcttgagat	cctttttttc	tgccgctaag	ctgctgcttg	caaacaaaaa	5040
aaccaccgct	accagcgggt	gtttgtttgc	cggatcaaga	gctaccaact	ctttttccga	5100
aggtaactgg	cttcagcaga	gcgcagatac	caaataactgt	ccttctagtg	tagccgtagt	5160
taggccacca	cttcaagaac	tctgtagcac	cgcctacata	cctcgctctg	ctaactcctgt	5220
taccagtggc	tgctgccagt	ggcgataagt	cgtgtcttac	cgggttgga	tcaagacgat	5280
agttaccgga	taaggcgcag	cggctcgggt	gaacgggggg	ttcgtgcaca	cagcccagct	5340
tggagcgaac	gacctacacc	gaactgagat	acctacagcg	tgagctatga	gaaagcgcca	5400

```

cgcttcccga agggagaaag gcgacaggt atccggtaag cggcagggtc ggaacaggag 5460
agcgacagag ggagcttcca ggggaaacg cctggatatc ttatagtcct gtcgggtttc 5520
gccacctctg acctgagcgt cgatttttgt gatgctcgtc aggggggcg agcctatgga 5580
aaaacgccag caacgcggcc tttttacggt tcttgccctt ttgctggcct tttgctcaca 5640
tgttctttcc tgcgttatcc cctgattctg tggataaccg tattaccgcc tttgagttag 5700
ctgataccgc tcgccgcagc cgaacgaccg agcgacagca gtcagtgagc gaggaagcgg 5760
aagagcgctt gatgcggtat tttctcctta cgcactctgt cgggtatttc caccgcatat 5820
gggtgactct cagtacaatc tgcctctgat ccgcatagtt aagccagtat acactccgct 5880
atcgctacgt gactgggtca tggctgcgcc ccgacacccg ccaacacccg ctgacgcgcc 5940
ctgacgggct tgtctgctcc cggcatccgc ttacagacaa gctgtgaccg tctccgggag 6000
ctgcatgtgt cagaggtttt caccgtcatc accgaaacgc gcgaggcagc tgcggtaaag 6060
ctcatcagcg tggctcgtgaa gcgattcaca gatgtctgcc tgttcatccg cgtccagctc 6120
gttgagtttc tccagaagcg ttaatgtctg gcttctgata aagcgggcca tgttaagggc 6180
ggttttttcc tgtttgggtca cttgatgcct ccgtgtaagg ggggaatttct gttcatgggg 6240
gtaatgatac cgatgaaacg agagaggatg ctacagatac gggttactga tgatgaacat 6300
gcccggttac tggaaacgtt tgagggtaaa caactggcgg tatggatgag gcgggaccag 6360
agaaaaatca ctacgggtca atgccagcgc ttogttaata cagatgtagg tgttccacag 6420
ggtagccagc agcatcctgc gatgcagatc cggaaacataa tgggtgcaggg cgctgacttc 6480
cgcgtttcca gactttacga aacacggaaa ccgaagacca ttcatgttgt tgctcaggtc 6540
gcagacgttt tgcagcagca gtcgcttcac gttcgctcgc gtatcgggtga ttcatctgc 6600
taaccagtaa ggcaaccccg ccagcctagc cgggtcctca acgacaggag cacgatcatg 6660
cgcacccgtg gccaggaccc aacgctgccg gagatgcgcc gcgtgcggct gctggagatg 6720
gcggacgcga tggatatgtt ctgccaaggg ttgggttgcg cattcacagt tctccgcaag 6780
aattgattgg ctccaattct tggagtgggt aatccgttag cgaggtgccg ccggcttcca 6840
ttcaggtcga ggtggcccg ctccatgcac cgcgacgcaa cgcggggagg cagacaaggt 6900
atagggcggc gcctacaatc catgccaaac cgttccatgt gctcgccgag gcggcataaa 6960
tcgccgtgac gatcagcggg ccagtgatcg aagttaggct ggtaagagcc gcgagcgatc 7020
cttgaagctg tccctgatgg tcgtcatcta cctgcctgga cagcatggcc tgcaacgcgg 7080
gcatcccgat gccgcgggaa gcgagaagaa tcataatggg gaaggccatc cagcctcgcg 7140

```

<210> 22

<211> 6725

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: vector

<220>

<223> pAN636

<400> 22

```

tcggcgcccg cttcgctgac cgaaacagca gttataaggc atgaagctgt ccggtttttg 60
caaaagtggc tgtgactgta aaaagaaatc gaaaaagacc gttttgtgtg aaaacggtct 120
ttttgtttcc ttttaaccaa ctgccataac tcgaggccta cctagcttcc aagaaagata 180
tcctaacagc acaagagcgg aaagatgttt tgttctacat ccagaacaac ctctgctaaa 240
attcctgaaa aattttgcaa aaagtgtgtg actttatcta caaggtgtgg tataataatc 300
ttaacaacag caggacgctc tagaaggagg agaaaacatg aaaacaaaac tggattttct 360
aaaaatgaag gagtctgaag aaccgattgt catgctgacc gcttatgatt atccggcagc 420
taaacttgct gaacaagcgg gagttgacat gatttttagtc ggtgattcac ttggaatgg 480
cgtcctcggc cttgattcaa ctgtcggtgt gacagttgag gatattgacc atcatacaa 540
agccgttaaa aggggtgcgc cgaatacctt tattgtgaca gatatgcggt ttatgtctta 600
tcacctgtct aaggaagata cgctgaaaaa tgcagcggct atcgttcagg aaagcggagc 660
tgacgcactg aagcttgagg gcggagaagg cgtgtttgaa tccattcgcg cattgacgct 720
tggaggcatt ccagtagtca gtcacttagg tttgacaccg cagtcagtcg gcgtactggg 780
cggctataaa gtacagggca aagacgaaca aagcgccaaa aaattaatag aagacagtat 840
aaaatgcgaa gaagcaggag ctatgatgct tgtgctggaa tgtgtgccgg cagaactcac 900
agccaaaatt gccgagacgc taagcatacc ggtcattgga atcggggctg gtgtgaaagc 960
ggacggacaa gttctcggtt atcatgatat tatcgccac ggtgttgaga gaacacctaa 1020
atttgtaaaag caatatacgc gcattgatga aaccatcgaa acagcaatca gcggatatgt 1080

```

tcaggatgta	agacatcgtg	ctttccctga	acaaaagcat	tcctttcaaa	tgaaccagac	1140
agtgccttgac	ggccttgtag	ggggaaaata	agggggggat	cctgttttgg	cggatgagag	1200
aagattttca	gcctgataca	gattaaatca	gaacgcagaa	gcggtctgat	aaaacagaat	1260
ttgcctggcg	gcagtagcgc	ggtggtccca	cctgacccca	tgccgaactc	agaagtgaat	1320
cgccgtagcg	ccgatggtag	tgtgggtct	ccccatgcga	gagtagggaa	ctgccaggca	1380
tcaaataaaa	cgaaaggctc	agtcgaaaga	ctgggccttt	cgttttatct	gttgtttgtc	1440
ggtgaacgct	ctcctgagta	ggacaaatcc	gccgggagcg	gatttgaacg	ttgcgaagca	1500
acggcccgga	gggtggcggg	caggacgccc	gccataaact	gccaggcatc	aaattaagca	1560
gaaggccatc	ctgacggatg	gcctttttgc	gtttctacaa	actcttggtg	ccgagacgat	1620
cgtcctcttt	gtttagtagcc	atcacttttg	ctgaagagta	ggagccgaaa	gtgacggcgt	1680
attcattgag	cggcagctga	gtcgcaccga	cagaaatcgc	ttctcttgat	gtgccggcg	1740
atccgactgt	ccagccgttc	ggtccgctgt	tgccgtttga	ggtaacagcg	acaacgcctt	1800
ctgacatggc	ccagtcaagc	gctgtgcttg	tcgcccagtc	cgggttggtt	aaagagtttc	1860
cgagagacag	gttcatcaca	tctgccccgt	cctgcactgc	acgttccacg	cccgcgatga	1920
cgttttccgt	tgtgccgctt	ccgccaggcc	ctaacacacg	ataagcaaga	agtgtggcat	1980
caggcgctac	gcctttaatc	gttccgtttg	cagccacagt	tccggctacg	tgtgtgccat	2040
ggtcagttgc	ctcgccccctc	ggatcgccgg	ttggtgtttc	ttttggatcg	taatcattgt	2100
ccacaaaatc	gtatccttta	tattgtccaa	agtttttctt	cagatctggg	tgattgtatt	2160
caaccccgat	gtcaataatc	gccacettga	tgctttttcc	tgtgtagcct	aatcccatg	2220
catcgtttgc	tccgatataa	ggcgcaactgt	catccatttg	cggagatacg	gcgtcttcgg	2280
agattgtggg	gaattctcat	gtttgacagc	ttatcatgca	atagttaccc	ttattatcaa	2340
gataagaaag	aaaaggattt	ttcgctacgc	tcaaatcctt	taaaaaaaca	caaaagacca	2400
cattttttaa	tgtggtcttt	attcttcaac	taaagcacc	attagttaa	caaacgaaaa	2460
ttggataaag	tgggataattt	ttaaaatata	tatttatgtt	acagtaatat	tgacttttaa	2520
aaaaggattg	attctaataga	agaaagcaga	caagtaagcc	tcctaaattc	actttagata	2580
aaaatttagg	aggcatatca	aatgaacttt	aataaaattg	atttagacaa	ttggaagaga	2640
aaagagatat	ttaatcatta	tttgaaccaa	caaacgactt	ttagtataac	cacagaaatt	2700
gatattagtg	ttttataaccg	aaacataaaa	caagaaggat	ataaatttta	ccctgcattt	2760
attttcttag	tgacaagggg	gataaaactca	aatacagctt	ttagaactgg	ttacaatagc	2820
gacggagagt	taggttattg	ggataagtta	gagccacttt	atacaatttt	tgatgggtga	2880
tctaaaacat	tctctggtat	ttggactcct	gtaaagaatg	acttcaaaga	gttttatgat	2940
ttataccttt	ctgatgtaga	gaaatataat	ggttcgggga	aattgtttcc	caaaacacct	3000
atacctgaaa	atgctttttc	tctttctatt	attccatgga	cttcattttac	tgggtttaac	3060
ttaaatatca	ataataatag	taattacctt	ctaccattta	ttacagcagg	aaaattcatt	3120
aataaaggta	attcaatata	tttaccgcta	tctttacagg	tacatcattc	tgtttgat	3180
ggttatcatg	caggattggt	tatgaactct	attcaggaat	tgtcagatag	gcctaattgac	3240
tggtttttat	aatatgagat	aatgccgact	gtacttttta	cagtccgttt	tctaattgtca	3300
ctaacctgcc	ccgttagttg	aagaacgaag	cggccgcaat	tcttgaagac	gaaagggcct	3360
cgtgatacgc	ctatttttat	aggttaatgt	catgataata	atggtttctt	agacgtcagg	3420
tggcactttt	cggggaaatg	tgcgcggaac	ccctattttg	ttatttttct	aaatacattc	3480
aaatatgtat	ccgctcatga	gacaataacc	ctgataaatg	cttcaataat	attgaaaaag	3540
gaagagtatg	agtattcaac	atttccgtgt	cgccttatt	cccttttttg	cggcattttg	3600
ccttccgtgt	tttgctcacc	cagaaacgct	ggtgaaagta	aaagatgctg	aagatcagtt	3660
gggtgccaga	gtgggttaca	tcgaactgga	tctcaacagc	ggtaagatcc	ttgagagttt	3720
tcgccccgaa	gaacgttttc	caatgatgag	cacttttaaa	gttctgctat	gtggcgcggt	3780
attatcccgt	attgacgccg	ggcaagagca	actcggtcgc	cgcatacact	attctcagaa	3840
tgacttggtt	gagtactcac	cagtcacaga	aaagcatctt	acggatggca	tgacagtaag	3900
agaattatgc	agtgtgccca	taacctagag	tgataacact	gcggccaact	tacttctgac	3960
aacgatcgga	ggaccgaagg	agctaaccgc	ttttttgcac	aacatggggg	atcatgtaac	4020
tcgccttgat	cgttgggaac	cggagctgaa	tgaagccata	ccaaacgacg	agcgtgacac	4080
cacgatgcct	gcagcaatgg	caacaacggt	gcgcaaaacta	ttaactggcg	aactacttac	4140
tctagcttcc	cggcaacaat	taatagactg	gatggaggcg	gataaagttg	caggaccact	4200
tctgcgctcg	gccttccgg	ctggctggtt	tattgctgat	aaatctggag	ccggtgagcg	4260
tgggtctcgc	ggtatcattg	cagcactggg	gccagatggt	aagccctccc	gtatcgtagt	4320
tatctacacg	acggggagtc	aggcaactat	ggatgaacga	aatagacaga	tcgctgagat	4380
aggtgcctca	ctgattaagc	attggtaact	gtcagaccaa	gtttactcat	atatacttta	4440
gattgattta	aaacttcatt	tttaatttaa	aaggatctag	gtgaagatcc	tttttgataa	4500
tctcatgacc	aaaatccctt	aacgtgagtt	ttcgttccac	tgagcgtcag	accccgtaga	4560
aaagatcaaa	ggatcttctt	gagatccttt	ttttctgcgc	gtaatctgct	gcttgcaaac	4620
aaaaaaacca	ccgctaccag	cgggtggtttg	tttgccggat	caagagctac	caactctttt	4680

```

tccgaaggta actggcttca gcagagcgca gataccaaat actgtccttc tagtgtagcc 4740
gtagttaggc caccacttca agaactctgt agcaccgcct acatacctcg ctctgctaata 4800
cctgttacca gtggctgctg ccagtggcga taagtcgtgt cttaccgggt tggactcaag 4860
acgatagtta ccgataaagg cgcagcggtc gggctgaacg ggggggttcgt gcacacagcc 4920
cagcttgag cgaacgacct acaccgaact gagataccta cagcgtgagc tatgagaaag 4980
cgccacgctt cccgaaggga gaaaggcgga caggatatccg gtaagcggca gggtcggaac 5040
aggagagcgc acgaggggagc ttccaggggg aaacgcctgg tatctttata gtcctgtcgg 5100
gtttcgccac ctctgacttg agcgtcgatt tttgtgatgc tcgtcagggg ggcggagcct 5160
atggaaaaac gccagcaacg cggccttttt acggttcctg gccttttgct ggccttttgc 5220
tcacatgttc tttcctgcgt tatccctga ttctgtggat aaccgtatta ccgcctttga 5280
gtgagctgat accgctcgcc gcagccgaac gaccgagcgc agcagtcag tgagcgagga 5340
agcgggaagag cgctgatgc ggtattttct ccttacgcac ctgtgcggta tttcacaccg 5400
catatgggtgc actctcagta caatctgctc tgatgccgca tagttaagcc agtatacaact 5460
ccgctatcgc tacgtgactg ggtcatggct gcgccccgac acccgccaac acccgctgac 5520
gcgcctgac gggcttgtct gctcccgga tccgcttaca gacaagctgt gaccgtctcc 5580
gggagctgca tgtgtcagag gttttcaccg tcatcaccga aacgcgcgag gcagctgcgg 5640
taaagctcat cagcgtggtc gtgaagcgat tcacagatgt ctgcctgttc atccgcgtcc 5700
agctcgttga gtttctccag aagcgttaat gtctggcttc tgataaagcg ggccatgtta 5760
agggcggttt tttcctgttt ggtcacttga tgccctcgtg taagggggaa tttctgttca 5820
tgggggtaat gataccgatg aaacgagaga ggatgtcac gatacgggtt actgatgatg 5880
aacatgcccg gttactggaa cgttgtgagg gtaacaact ggcgggtatg atgcggcggg 5940
accagagaaa aatcactcag ggtcaatgcc agcgttcgt taatacagat gtaggtgttc 6000
cacagggtag ccagcagcat cctgcgatgc agatccggaa cataatggtg cagggcgctg 6060
acttcgcgt ttcagactt tacgaaacac ggaaaccgaa gaccattcat gttgttgctc 6120
aggtcgcaga cgttttgcag cagcagtcgc ttcacgttcg ctgcgtatc ggtgattcat 6180
tctgctaacc agtaaggcaa cccgcgcag ctagcgggt cctcaacgac aggagcacga 6240
tcatgcgcac ccgtggccag gaccacaacgc tgcgcgagat gcgcgcgtg cggctgctgg 6300
agatggcgga cgcgatggat atgttctgcc aagggttggt ttgcgcattc acagttctcc 6360
gcaagaattg attggtcca attcttgagg tggatgaatcc gttagcgagg tgccgcccggc 6420
ttccattcag gtcgaggtgg cccggctcca tgaccgcga cgcaacgcgg ggaggcgagc 6480
aaggtatagg gcggcgccca caatccatgc caaccgctc catgtgctcg ccgaggcggc 6540
ataaatcgcc gtgacgatca gcggtccagt gatcgaagtt aggctggtaa gagccgcgag 6600
cgatccttga agctgtccct gatggctcgtc atctacctgc ctggacagca tggcctgcaa 6660
cgcgggcacc ccgatgccgc cggaagcgag aagaatcata atggggaagg ccatccagcc 6720
tcgcg 6725

```

<210> 23

<211> 6806

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: vector

<220>

<223> pAN637

<400> 23

```

tcggcgggccg cttcgtcgac cgaaacagca gttataaggc atgaagctgt ccgggtttttg 60
caaaagtggc tgtgactgta aaaagaaatc gaaaaagacc gttttgtgtg aaaacggtct 120
ttttgtttcc ttttaaccaa ctgccataac tcgaggccta cctagcttcc aagaaagata 180
tcctaacagc acaagagcgg aaagatgttt tgttctacat ccagaacaac ctctgctaaa 240
attcctgaaa aattttgcaa aaagtgtttg actttatcta caaggtgtgg tataataatc 300
ttaacaacag caggacgctc tagacaattg agatcttaag aaaggaggtg ttaattaatg 360
aagattggaa tcattggcgg aggtccggtt ggtcttttat gcgcctatta tttgtcactt 420
tatcacgacg tgactgttgt gacgaggcgg caagaacagg ctgcggccat tcagtctgaa 480
ggaatccggc tttataaagg cggggaggaa ttcagggtcg attgcagtgc ggacacgagt 540
atcaattcgg actttgacct gcttgctcgtg acagtgaagc agcatcagct tcaatctgtt 600
ttttcgtcgc ttgaacgaat cggaagacg aatatattat ttttgcaaaa cggcatgggg 660
catatccacg acctaaaaga ctggcacgtt ggccattcca tttatgttgg aatcggttag 720

```


cacggagctg	taagaaaatc	ggatacagct	gttgatcata	caggcctagg	tgcgataaaa	780
tggagcgcgt	tcgacgatgc	tgaaccagac	cggctgaaca	tcttgtttca	gcataaccat	840
tcggattttc	cgatttatta	tgagacggat	tggtagcgtc	tgctgacggg	caagctgatt	900
gtaaatgcgt	gtattaatcc	tttaactgcg	ttattgcaag	tgaaaaatgg	agaactgctg	960
acaacgccag	cttatctggc	ttttatgaag	ctggattttc	aggaggcatg	ccgcatttta	1020
aaacttgaaa	atgaagaaaa	ggcttgggag	cgggttcagg	ccgtttgtgg	gcaaacgaaa	1080
gagaatcggt	catcaatgct	ggttgacgtc	attggaggcc	ggcagacgga	agctgacgcc	1140
attatcggat	acttattgaa	ggaagcaagt	cttcaaggct	ttgatgccgt	ccacctagag	1200
tttttatatg	gcagcatcaa	agcattggag	cgaatatcca	acaaagtggg	ttactaagga	1260
tcctgttttg	gcggatgaga	gaagattttc	agcctgatac	agattaaatc	agaacgcaga	1320
agcggctctg	taaaacagaa	tttgctggc	ggcagtagcg	cgggtggccc	acctgacccc	1380
atgccgaact	cagaagtga	acgccgtagc	gccgatggta	gtgtggggtc	tcccatgctg	1440
agagtaggga	actgccaggc	atcaaataaa	acgaaaggct	cagtcgaaag	actgggcctt	1500
tcgttttatc	tgttgtttgt	cgggtgaacgc	tctcctgagt	aggacaaatc	cgccggggagc	1560
ggatttgaac	gttgcggaagc	aacggcccgg	aggggtggcg	gcaggacgcc	cgccataaac	1620
tgccaggcat	caaattaagc	agaaggccat	cctgacggat	ggcctttttg	cgtttctaca	1680
aactcttggt	accgagacga	tcgtcctctt	tgttgtagcc	catcactttt	gctgaagagt	1740
aggagccgaa	agtgacggcg	tattcattga	gcggcagctg	agtcgcaccg	acagaaatcg	1800
cttctcttga	tgtgcccggc	gatccgactg	tccagccggt	cggctccgctg	ttgccgtttg	1860
aggtaacagc	gacaacgcct	tctgacatgg	cccagtcaag	cgtgtgctt	gtcgcccagt	1920
ccgggttggt	taaagagttt	ccgagagaca	ggttcacac	atctgccccg	tcctgcaactg	1980
caegtccac	gcccgcgatg	acgttttcog	ttgtgcogct	tccgccaggc	cctaacacac	2040
gataagcaag	aagtgtggca	tcaggcgcta	cgcctttaat	cgttccggtt	gcagccacag	2100
ttccggctac	gtgtgtgcca	tggtcagttg	cctgcgccct	cggatcgccg	gttggtgttt	2160
cttttggtac	gtaatcattg	tccacaaaat	cgtatccttt	atattgtcca	aagtttttct	2220
tcagatctgg	gtgattgtat	tcaaccccag	tgtcaataat	cgccaccttg	atgccttttc	2280
ctgtgtagcc	taaatcccat	gcacgttttg	ctccgatata	aggcgcaactg	tcatccattt	2340
gcggagatac	ggcgtcttcg	gagattgtgg	ggaattctca	tgtttgacag	cttatcatgc	2400
aatagttacc	cttattatca	agataagaaa	gaaaaggatt	tttcgctacg	ctcaaactct	2460
ttaaaaaac	acaaaagacc	acatttttta	atgtggtctt	tattcttcaa	ctaaagcacc	2520
cattagttca	acaaacgaaa	attggataaa	gtgggatatt	tttaaaatat	atatttatgt	2580
tacagtaata	ttgactttta	aaaaaggatt	gattctaattg	aagaaagcag	acaagtaagc	2640
ctcctaaatt	cacttttagat	aaaaatttag	gaggcatatc	aatgaactt	taataaaatt	2700
gatttagaca	attggaagag	aaaagagata	tttaatcatt	atttgaacca	acaaacgact	2760
tttagtataa	ccacagaaat	tgatattagt	gttttataacc	gaaacataaa	acaagaagga	2820
tataaatatt	accctgcatt	tattttctta	gtgacaaggg	tgataaaactc	aaatacagct	2880
tttagaactg	gttacaatag	cgacggagag	ttaggttatt	gggataagtt	agagccactt	2940
tatacaattt	ttgatggtgt	atctaaaaca	ttctctggta	tttggaactcc	tgtaaagaat	3000
gacttcaaag	agttttatga	tttatacctt	tctgatgtag	agaaatataa	tggttcgggg	3060
aaattgtttc	ccaaaacacc	tatacctgaa	aatgcttttt	ctctttctat	tattccatgg	3120
acttcattta	ctgggtttta	cttaaatatc	aataataata	gtaattacct	tctacctt	3180
attacagcag	gaaaattcat	taataaagg	aattcaatat	atttaccgct	atctttacag	3240
gtacatcatt	ctgtttgtga	tggttatcat	gcaggattgt	ttatgaactc	tattcaggaa	3300
ttgtcagata	ggcctaata	ctggctttta	taatattgaga	taatgccgac	tgtacttttt	3360
acagtcgggt	ttctaattgtc	actaacctgc	cccgttagtt	gaagaacgaa	gcggccgcaa	3420
ttcttgaaga	cgaaagggcc	tcgtgatacg	cctattttta	taggttaatg	tcatgataat	3480
aatggtttct	tagacgtcag	gtggcacttt	tcggggaaat	gtgcgcggaa	cccctatttg	3540
tttatttttc	taaatacatt	caaatatgta	tccgctcatg	agacaataac	cctgataaat	3600
gcttcaataa	tattgaaaaa	ggaagagtat	gagtattcaa	catttccgtg	tcgcccttat	3660
tccctttttt	gcggcatttt	gccttctctg	ttttgctcac	ccagaaacgc	tggtgaaagt	3720
aaaagatgct	gaagatcagt	tgggtgcacg	agtgggttac	atcgaactgg	atctcaacag	3780
cggtaagatc	cttgagagtt	ttcgccccga	agaacgtttt	ccaatgatga	gcacttttaa	3840
agttctgcta	tgtggcgcg	tattatcccg	tattgacgcc	gggcaagagc	aactcggctg	3900
ccgcatacac	tattctcaga	atgacttggt	tgagtactca	ccagtcacag	aaaagcatct	3960
tacggatggc	atgacagtaa	gagaattatg	cagtgtctgcc	ataaccatga	gtgataaacac	4020
tgccggccaac	ttacttctga	caacgatcgg	aggaccgaag	gagctaaccg	cttttttgca	4080
caacatgggg	gatcatgtaa	ctcgccttga	tcgttgggaa	ccggagctga	atgaagccat	4140
accaaacgac	gagcgtgaca	ccacgatgcc	tcgagcaatg	gcaacaacgt	tgcgcaaact	4200
attaactggc	gaactactta	ctctagcttc	ccggcaacaa	ttaatagact	ggatggaggc	4260
ggataaagtt	gcaggaccac	ttctgcgctc	ggcccttccg	gctggctggt	ttattgctga	4320

```

taaatctgga gccggtgagc gtgggtctcg cggtatcatt gcagcactgg ggccagatgg 4380
taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg 4440
aaatagacag atcgctgaga taggtgcctc actgattaag cattggtaac tgtcagacca 4500
agtttactca tatatacttt agattgattt aaaacttcat ttttaattta aaaggatcta 4560
ggtgaagatc ctttttgata atctcatgac caaaatccct taacgtgagt tttcgttcca 4620
ctgagcgtca gaccccgtag aaaagatcaa aggatcttct tgagatcctt tttttctgcg 4680
cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca gcggtggttt gtttgccgga 4740
tcaagagcta ccaactcttt ttccgaaggt aactggcttc agcagagcgc agataccaaa 4800
tactgtcctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc 4860
tacatacctc gctctgctaa tctgttacc agtggctgct gccagtggcg ataagtctgt 4920
tcttaccggg ttggactcaa gacgatagtt accggataag gcgcagcggg cgggctgaac 4980
gggggggttcg tgcacacagc ccagcttgga gcgaacgacc tacaccgaac tgagatacct 5040
acagcgtgag ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc 5100
ggtaagcggc agggctcgaa caggagagcg cacgaggag cttccagggg gaaacgcctg 5160
gtatctttat agtctgtctg ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg 5220
ctcgtcaggg gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttccct 5280
ggccttttgc tggccttttg ctacatgtt ctttctgctg ttatcccctg attctgtgga 5340
taaccgtatt accgcctttg agtgagctga taccgctcgc cgcagccgaa cgaccgagcg 5400
cagcgagtca gtgagcgagg aagcggaaga gcgcctgatg cggtattttc tccttacgca 5460
tctgtgcggg atttcacacc gcatatggtg cactctcagt acaatctgct ctgatgccgc 5520
atagttaagc cagtatacac tccgctatcg ctacgtgact gggctcatggc tgcgccccga 5580
caccgcca caccgctga cgcgccctga cgggcttgct tgctcccggc atccgcttac 5640
agacaagctg tgaccgtctc cgggagctgc atgtgtcaga ggttttcacc gtcacaccg 5700
aaacgcgcga ggcagctgcg gtaaaagtca tcagcgtggt cgtgaagcga ttcacagatg 5760
tctgctgtgt catccgcgtc cagctcgttg agtttctcca gaagcgttaa tgtctggtt 5820
ctgataaagc gggccatggt aaggcggtt ttttctgtt tggctacttg atgcctccgt 5880
gtaaggggga atttctgttc atgggggtaa tgataccgat gaaacgagag aggatgctca 5940
cgatacgggt tactgatgat gaacatgccc ggttactgga acgttgtgag ggtaaacaa 6000
tgccggtatg gatgcggcgg gaccagagaa aaatcactca gggatcaatgc cagcgcttcg 6060
ttaatacaga tgtaggtgtt ccacagggtg gccagcagca tcttgctgat cagatccgga 6120
acataatggt gcagggcgct gacttccgcg tttccagact ttacgaaaca cggaaaccga 6180
agaccattca tgttggttgc caggtcgcag acgttttgca gcagcagtcg cttcacgttc 6240
gctcgcgtat cgggtgattca ttctgctaac cagtaaggca accccgccag cctagccggg 6300
tctcaacga caggagcacg atcatgcgca cccgtggcca ggacccaacg ctgcccgaga 6360
tgccgcgcgt gcggctgctg gagatggcgg acgcgatgga tatgttctgc caagggttg 6420
tttgccgcat cacagttctc cgcaagaatt gattggctcc aattcttgga gtggtgaatc 6480
cgttagcgag gtgcccgcgg cttccattca ggtcagagtg gcccggtcc atgcaccgcg 6540
acgcaacgcg gggaggcaga caaggtatag ggcggcgcct acaatccatg ccaaccggtt 6600
ccatgtgctc gccgaggcgg cataaatcgc cgtgacgac agcgggtccag tgatcgaagt 6660
taggctggta agagccgcga gcgatccttg aagctgtccc tgatggtcgt catctacctg 6720
cctggacagc atggcctgca acgcgggcat cccgatgccg ccggaagcga gaagaatcat 6780
aatggggaag gccatccagc ctccgcg 6806

```

<210> 24

<211> 3867

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: vector

<400> 24

```

aagctttctc aagaagcga caagaaaaaa gaagagcaga ttaaacagct tcaagagttt 60
gtcgtagat tcagcgccaa tgcgtctaaa tctaagcagg ctacatcaag aaagaaactt 120
ctcgaaaaaa tcacgctgga tgatattaaa ccgtcttccc gccgctatcc ttatgttaac 180
ttcacgcggg aacgggaaat cggaaatgat gttcttcgcy tggaaggctt acaaaaaaca 240
atcgatggcg taaaggtgct tgacaatgtc agctttatta tgaatcgaga agataaaatt 300
gctttcactg gccgaaatga acttgctgtt actacgctgt ttaaaatcat ttccggggaa 360
atggaagctg acagcgggaa gtttaaatgg ggtgttacca catctcaagc gtattttcca 420
aaagacaaca gcgaatattt cgaaggcagt gatctgaacc ttgtagactg gcttcgcca 480

```


tattctccgc	acgaccaaag	tgagagcttt	ttacgcgggt	tcttaggacg	catgctgttc	540
tctggagaag	aagtccacaa	aaaagcaaat	gtactttccg	ggggagaaaa	ggtccgctgt	600
atgctgtcga	aagcgatgct	ttctggcgcc	aatattttaa	ttttggatga	gccgaccaac	660
catttagacc	tagagtcctat	tacagcgctc	aataacggct	taatcagctt	taaaggcgct	720
atgctgttta	cttcccatga	ccatcagctt	gtgcagacca	ttgccaacag	aattatagaa	780
attacaccta	acggcatcgt	cgataagcaa	atgagctatg	acgagttcct	tgaaaatgct	840
gatgtgcaga	aaaaattgac	tgaactatac	gccgaataaa	aaagcagaga	tttctctgct	900
ttttttgata	cctaaatgtg	aaaggagatc	acaacatgaa	atttttgggt	gtcggagcag	960
gtggagtagg	cgggtatatt	ggcggacggc	tttcggagaa	aggaaatgat	gtgacatttc	1020
tcgtgcgcca	aaaacgagct	gagcagctta	aaaaaacggg	gcttgtcatc	catagtgaag	1080
aagggaaatgt	atcatttcag	cccgaactaa	tcagtgcggg	agaaacaggg	caatttgatg	1140
tcgttatcat	tgcttctaaa	gcatactcgc	ttggtcaagt	gatagaggat	gtcaaaccat	1200
ttatccatca	agaatctgtc	attatccctt	ttttaaatgg	gtaccgccac	tatgagcagc	1260
tatttgccgc	attttcaaaa	gaacaggtgc	tgggcccgtt	gtgttttata	gaaagtgttt	1320
tagacaacaa	aggagaaatt	catcatacga	gcgcacgcga	tcgtttttgt	tttgagagaa	1380
ggaacggcga	gcgtacggag	cggataagag	cgcttgaaga	ggcattttca	ggtgtgaagg	1440
ctgaagtcat	cattagcggg	catatcgaga	agatcccctg	cagcaatagt	tacccttatt	1500
atcaagataa	gaaagaaaaa	gatttttctc	tacgctcaaa	tccttttaaa	aaacacaaaa	1560
gaccacattt	tttaattgtg	tctttattct	tcaactaaag	cacccattag	ttcaacaaac	1620
gaaaattgga	taaagtggga	tattttttaa	atataatatt	atgttacagt	aatattgact	1680
tttaaaaaag	gattgattct	aatgaagaaa	gcagacaagt	aagcctccta	aattcacttt	1740
agataaaaaa	ttaggaggca	tatcaaataa	actttaataa	aattgattta	gacaattgga	1800
agagaaaaga	gatatttaaa	cattatttga	accaacaaac	gacttttagt	ataaccacag	1860
aaattgatat	tagtggttta	taccgaaaca	taaaacaaga	aggatataaa	ttttaccctg	1920
catttatttt	cttagtgaca	aggggtgata	actcaaatac	agctttttag	actggttaca	1980
atagcgacgg	agagttaggt	tattgggata	agttagagcc	actttatata	atttttgatg	2040
gtgtatctaa	aacattctct	ggtatttgga	ctcctgtaaa	gaatgacttc	aaagagtttt	2100
atgattttata	cctttctgat	gtagagaaat	ataatggttc	ggggaaattg	tttcccaaaa	2160
cacctatacc	tgaaaatgct	ttttctcttt	ctattattcc	atggacttca	tttactgggt	2220
tttaacttaa	tatcaataat	aatagtaatt	accttctacc	cattattaca	gcaggaaaat	2280
tcattaataa	aggttaattca	atataattac	cgctatcttt	acagggtacat	cattctgttt	2340
gtgatgggta	tcagtcagga	ttgtttatga	actctattca	ggaattgtca	gataggccta	2400
atgactggct	tttataatat	gagataatgc	cgactgtact	ttttacagtc	ggttttctaa	2460
tgtcactaac	ctgccccgtt	agttgaagaa	ggtttttata	ttacagctcc	cgggagatct	2520
gggatcacta	gtccaaacga	cagaaggcga	ccacctgcat	ggatttttga	ttgaaaaagc	2580
aaaacgttta	tctctcgctg	caccagtatt	agaaaccgtt	tatgcgaatc	tgcaaatgta	2640
tgaagcagaa	aaataaaaaa	aggaggcgga	aaagcctcct	tttattttact	taaaaagccc	2700
aatttccggt	tctgaagata	ggctctcttt	tcccgctctg	cgtaattccg	tcaatattca	2760
tatccttaga	accgatcata	aagtccacgt	gtgtaatgct	ttcatttagg	ccttctttga	2820
caagctcttc	acgagacatc	tgctttccgc	cttcaatatt	aaaggcatag	gcgcttccga	2880
tcgccaaatg	atgttgacgcg	ttttcatcaa	acagcgtggt	atagaaaaga	atgtttgatt	2940
gtgatataag	cgaatcgtaa	ggaacaagtg	ccacttcacc	taaatagtga	gaaccttcac	3000
ctgtttccac	cagttctttt	aaaatatcct	cacttttttc	agctttaatg	tcgactatac	3060
ggccattttc	aaacgtcagg	gtgaaatttt	caataatatt	tccgcgtag	cttaattggt	3120
ttgtgcttga	taccactccg	tcaaccccg	ctttttgcgg	cagcgtgaac	acttcttctg	3180
tcggcatatt	ggccataaac	tcattggccac	tttcattcac	gcttcccgca	cctgccccaa	3240
catgcttcct	aggcagctta	attgttagat	cagttccctc	tgcttgataa	tgtaaggcag	3300
cgtaatgtct	ctcggtcaaa	tggtcaactt	tttcatgaag	attttgggtc	tgattgatcc	3360
acgctgaac	cgggttgtct	tcatttacgc	gcgtcgcttt	aaaaatttct	tcccacagaa	3420
ggtggatcgc	ttctcctctc	gatttgccag	gaaacacctt	gtgagcccag	cctgctgatg	3480
ccgcacctac	gacagtccag	ctgactttgt	ctgattgaat	atattgtctg	tatgtatgta	3540
atgctttgcc	tgctgctttt	tggaatgccg	caatccgttt	ggaatctata	ccttttagca	3600
agtctgggtt	cgacgacaca	acagaaaatga	aagcagctcc	atttttggca	agctcttctc	3660
tgccttttgc	ttcccattca	ggatattctt	caaagtcttc	aaacggcgca	agttcgtatt	3720
tttaatttggc	gacttcgtca	tcttgccaat	tcacgggtgac	gttttttgcg	cccttttcat	3780
atgctgtgtt	tacaattaaa	cggacaaaat	cccgaacgtc	tgttgaagca	tttacgacta	3840
catactggcc	tttttggaca	tttaacgc				3867

<210> 25
 <211> 8704
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: vector

<220>
 <223> pAN238

<400> 25
 gcgggccgctt cgtcgaccga aacagcagtt ataaggcatg aagctgtccg gtttttgcaa 60
 aagtggctgt gactgtaaaa agaaatcgaa aaagaccgtt ttgtgtgaaa acgggtctttt 120
 tgtttccttt taaccaactg ccataactcg aggcctacct agcttccaag aaagatatcc 180
 taacagcaca agagcggaaa gatgttttgt tctacatcca gaacaacctc tgctaaaatt 240
 cctgaaaaat tttgcaaaaa gttgttgact ttatctacaa ggtgtggtat aataatctta 300
 acaacagcag gacgctctag aggaggagac taacatgaaa tttttgggtg tcggagcagg 360
 tggagtaggc gggatatatt ggcggacggc ttccggagaaa ggaaatgatg tgacatttct 420
 cgtgcgccaa aaacgagctg agcagcttaa aaaaaccggg cttgtcatcc atagtgtaaa 480
 aggaatgta tcatttcagc ccgaactaat cagtgcggga gaaacagggc aatttgatgt 540
 cgttatcatt gcttctaaag catactcgct tggtaagtg atagaggatg tcaaacatt 600
 tatccatcaa gaatctgtca ttatcccttt tttaaattggg taccgccact atgagcagct 660
 atttgcggca ttttcaaaag aacaggtgct gggcggcctg tgttttatag aaagtgtctt 720
 agacaacaaa ggagaaattc atcatacgag cgcacgcgat cgttttgtat ttggagaatg 780
 gaacggcgag cgtacggagc ggataagagc gcttgaagag gcattttcag gtgtgaaggc 840
 tgaagtcata attagcgggc atatcgagaa ggacatttgg aaaaagtatc tctttattgc 900
 agcgcaagcg gggatcacia cgttatttca acgaccgctt ggcccaatcc tcgccacaga 960
 agcggagcgt cacacggccc aaactcttat tggggaaatt tgcaactgtt tacgaaaaga 1020
 aggtgttccg gctgatccgg ctcttgagga agagagcttt cgtacgatga ccagcatgtc 1080
 ttaccatatt aagtctctca tgcttcggga tatggaaaac ggccaaacga cagaaggcga 1140
 ccacctgcat ggatttttga ttgaaaaagc aaaacgttta tctctcgtg caccagtatt 1200
 agaaaccgtt tatgcgaatc tgcaaatgta tgaagcagaa aaataaaaaa aggaggcgga 1260
 aaagcctcct tttatttact taaaaagccc aatttccgtt tctgaagata ggctctcttt 1320
 tcccgtctgc cgggatacctg ttttgccgga tgagagaaga ttttcagcct gatacagatt 1380
 aaatcagaac gcagaagcgg tctgataaaa cagaatttgc ctggcggcag tagcgcgggtg 1440
 gtcccacctg acccatgccc gaactcagaa gtgaaacgcc gtagcgcgga tggtagtgtg 1500
 ggggtctccc atgcgagagt agggaaactg caggcatcaa ataaaacgaa aggtcagtc 1560
 gaaagactgg gcctttcgtt ttatctgttg ttgtcggtg aacgtctctc tgagtaggac 1620
 aaatccgccc ggagcggatt tgaacgttgc gaagcaacgg cccggagggt ggcgggcagg 1680
 acgcccgcga taaactgccg ggcatacaat taagcagaag gccatcctga cggatggcct 1740
 ttttgcgctt ctacaaactc ttggtaccca gaaaaagcgg caaaagcggc tgttaaaaaa 1800
 gcgaaatcga agaagctgtc tgccgctaag acggaatata aaaagcggtc tgctgttgtg 1860
 tcatctttaa aagtacagc cgaatgaatc cagaagatg tcctaaaata cttgaacacc 1920
 cagaaagata aaggaaatgc agaccaaatt cattcttatt atgtggtgaa cgggattgct 1980
 gttcatgcct caaaagaggt tatggaaaaa gtggtgcagt ttcccgaagt ggaaaagggtg 2040
 cttcctaatt agaaacggca gctttttaag tcatcctccc catttaatat gaaaaaagca 2100
 cagaaagcta ttaaagcaac tgacggtgtg gaatggaatg tagaccaaatt cgatgccccca 2160
 aaagcttggg cacttgagata tgatggaact ggcacgggtt ttgcgtccat tgataccggg 2220
 gtggaatgga atcatccggc attaaaagag aaatatcgcg gatataatcc ggaaaatcct 2280
 aatgagcctg aaaatgaaat gaactggtat gatgcgtag caggcgaggc aagcccttat 2340
 gatgatttgg ctcatggaac ccacgtgaca ggcacgatgg tgggctctga acctgatgga 2400
 acaaatcaaa tcgggtgtagc acctggcgca aaatggattg ctgttaaagc gttctctgaa 2460
 gatggcggca ctgatgctga cattttggaa gctggtgaat gggtttttagc accaaaggac 2520
 gcggaaggaa atccccaccc ggaaatggct cctgatgttg tcaataactc atggggaggg 2580
 ggctctggac ttgatgaatg gtacagagac atggtcaatg cctggcggtc ggccgatatt 2640
 ttccctgagt tttcagcggg gaatacggat ctctttatcc cggcggggccc tggttctatc 2700
 gcaaatccgg caaactatcc agaactggtt gcaactggag cgactgagaa ttccaattcc 2760
 ccatggagag aaaagaaaat cgctaattgt gattactttg aacttctgca tattcttgaa 2820
 tttaaaaagg ctgaaagagt aaaagattgt gctgaaatat tagagtataa acaaaatcgt 2880

```

gaaacaggcg aaagaaagtt gtatcgagtg tggttttgta aatccaggct ttgtccaatg 2940
tgcaactgga ggagagcaat gaaacatggc attcagtcac aaaaggttgt tgctgaagtt 3000
attaacaaca agccaacagt tcggttggttg tttctcacat taacagttaa aaatgtttat 3060
gatggcgaag aattaaataa gagtttgta gatatggctc aaggatttcg ccgaatgatg 3120
caatataaaa aaattaataa aaatcttggt ggttttatgc gtgcaacgga agtgacaata 3180
aataataaag ataattctta taatcagcac atgcatgtat tggatatgtg ggaaccaact 3240
tattttaaga atacagaaaa ctacgtgaat caaaaacaat ggattcaatt ttggaaaaag 3300
gcaatgaaat tagactatga tccaaatgta aaagttcaaa tgattcgacc gaaaaataaa 3360
tataaatcgg atatacaatc ggcaattgac gaaactgcaa aatatcctgt aaaggatacg 3420
gattttatga ccgatgatga agaaaagaat ttgaaacggt tgtctgattt ggaggaaggt 3480
ttacaccgta aaaggttaat ctctatggt ggtttgttaa aagaaataca taaaaatta 3540
aaccttgatg acacagaaga aggcgatttg attcatacag atgatgacga aaaagccgat 3600
gaagatggat tttctattat tgcaatgtgg aattgggaac ggaaaaatta ttttattaaa 3660
gagtagttca acaaacgggc catattgttg tataagtgat gaaatactga atttaaaact 3720
tagtttataat gtggtaaaaat gttttaatca agtttaggag gaattaatta tgaagtgtaa 3780
tgaatgtaac agggttcaat taaaagaggg aagcgtatca ttaacctat aaactacgtc 3840
tgccctcatt attggagggt gaaatgtgaa tacatcctat tcacaatcga atttacgaca 3900
caaccaaatt ttaatttggt tttgcaattt atcttttttt agcgtattaa atgaaatggt 3960
tttgaacgtc tcattacctg atattgcaaa tgattttaat aaaccacctg cgagtacaaa 4020
ctgggtgaac acagccttta tgtaaacctt ttccattgga acagctgtat atggaaagct 4080
atctgatcaa ttaggcataa aaaggttact cctatttgga attataataa attgtttcgg 4140
gtcggtaatt gggtttggtg gccattcttt cttttcctta cttattatgg ctcgttttat 4200
tcaaggggct ggtgcagctg catttccagc actcgtaatg gttgtagttg cgcgctatat 4260
tccaaaggaa aataggggta aagcatttggt tcttattgga tcgatagtag ccatgggaga 4320
aggatcggt ccagcgattg gtggaatgat agccattat attcattggt cctatcttct 4380
actcttctt actataacaa ttatcactgt tccgtttctt atgaaattat taaagaaaga 4440
agtaaggata aaaggtcatt ttgatataa aggaattata ctaatgtctg taggcattgt 4500
attttttatg ttgtttacaa catcatatag catttctttt cttatcgtaa gcgtgctgtc 4560
attcctgata tttgtaaaa acatcaggaa agtaacagat ccttttggtg atcccggtat 4620
agggaataat atacctttta tgattggagt tctttgtggg ggaattatat ttggaacagt 4680
agcagggttt gtctctatgg ttcttatat gatgaaagat gttcaccagc taagtactgc 4740
cgaaatcgga agtgtaatta ttttccctgg aacaatgagt gtcattattt tcggctacat 4800
tggtgggata cttgttgata gaagagggtc tttatacgtg ttaaaccatg gagttacatt 4860
tctttctgtt agctttttta ctgcttctt tcttttagaa acaacatcat ggttcatgac 4920
aattataatc gtatttggtt taggtgggct ttogttcacc aaaacagtta tatcaacaat 4980
tgtttcaagt agcttgaaac agcaggaagc tgggtgctgga atgagtttgc ttaactttac 5040
cagcttttta tcagagggaa caggatttgc aattgtagggt ggtttattat ccataccctt 5100
acttgatcaa aggttggttac ctatggaagt tgatcagtc acttatctgt atagtaattt 5160
gttattactt ttttcaggaa tcattgtcat tagttggtg gttacctga atgtatataa 5220
acattctcaa agggattttct aaatcgttaa gggatcaact ttgggagaga gttcaaaatt 5280
gatccttttt ttataacagt tcgaagcggc cgcaattctt gaagacgaaa gggcctcgtg 5340
atacgccat ttttataggt taatgtcatg ataataatgg tttcttagac gtcagggtggc 5400
acttttcggg gaaatgtgcg cggaaaccct atttgtttat ttttctaaat acattcaaat 5460
atgtatccgc tcatgagaca ataaccctga taaatgcttc aataatattg aaaaaggaag 5520
agtatgagta ttcaacattt ccgtgtcgcc cttattccct tttttgcggc attttgcctt 5580
cctgtttttg ctacccaga aacgctggtg aaagtaaaag atgctgaaga tcagttgggt 5640
gcacgagtg gttacatcga actggatctc aacagcggta agatccttga gagttttcgc 5700
cccgaagaac gttttccaat gatgagcact tttaaagttc tgctatgtgg cgcggtatta 5760
tcccgatttg acgcgggca agagcaactc ggtcgcgca tacactattc tcagaatgac 5820
ttggttgagt actcaccagt cacagaaaag catcttacgg atggcatgac agtaagagaa 5880
ttatgcagtg ctgccataac catgagtgat aacactgcgg ccaacttact tctgacaacg 5940
atcgaggag cgaaggagct aaccgctttt ttgcacaaca tgggggatca tgtaactcgc 6000
cttgatcggt gggaaccgga gctgaatgaa gccataccaa acgacgagcg tgacaccacg 6060
atgcctgcag caatggcaac aacgttgctc aaactattaa ctggcgaact acttactcta 6120
gcttcccggc aacaattaat agactggatg gaggcggata aagttgcagg accacttctg 6180
cgctcgccc ttccggctgg ctggtttatt gctgataaat ctggagccgg tgagcgtggg 6240
tctcgcggt tcatigcagc actggggcca gatggtgaag cctcccgat cgtagttatc 6300
tacacgacgg ggagtcaggc aactatggat gaacgaaata gacagatcgc tgagataggt 6360
gcctcactga ttaagcattg gtaactgtca gaccaagttt actcatatat actttagatt 6420
gatttaaaac ttcattttta atttaaaagg atctaggtga agatcctttt tgataatctc 6480

```

atgaccaaaa	tcccttaacg	tgagttttcg	ttccaactgag	cgtcagaccc	cgtagaaaaag	6540
atcaaaggat	cttcttgaga	tccttttttt	ctgcgcgtaa	tctgctgctt	gcaaacaaaa	6600
aaaccaccgc	taccagcggg	ggttttgttg	ccggatcaag	agctaccaac	tctttttccg	6660
aaggtaactg	gcttcagcag	agcgcagata	ccaaatactg	tccttctagt	gtagccgtag	6720
ttaggccacc	acttcaagaa	ctctgtagca	ccgcctacat	acctcgctct	gctaatacctg	6780
ttaccagtgg	ctgctgccag	tggcgataag	tcggtgttta	ccgggttggg	ctcaagacga	6840
tagttaccgg	ataaggcgca	gcggtcgggc	tgaacggggg	gttcgtgcac	acagcccagc	6900
ttggagcgaa	cgacctacac	cgaactgaga	tacctacagc	gtgagctatg	agaaagcgcc	6960
acgcttcccg	aaggagagaa	ggcggacagg	tatccggtaa	gcggcagggt	cggaacaggga	7020
gagcgcacga	gggagcttcc	agggggaaac	gcctggatc	tttatagtc	tgctcggttt	7080
cgccacctct	gacttgagcg	tcgatttttg	tgatgctcgt	caggggggcg	gagcctatgg	7140
aaaaacgcca	gcaacgcggc	ctttttacgg	ttcctggcct	tttgctggcc	ttttgctcac	7200
atgttctttc	ctgcgttatc	ccctgattct	gtggataacc	gtattaccgc	ctttgagtga	7260
gctgataccg	ctcgccgcag	ccgaacgacc	gagcgcagcg	agtcagttag	cgaggaagcg	7320
gaagagcgcc	tgatgcggta	ttttctcctt	acgcattctgt	gcggtatttc	acaccgcata	7380
tggtgcactc	tcagtacaat	ctgctctgat	gccgcatagt	taagccagta	tacactccgc	7440
tatcgctacg	tgactgggtc	atggctgcgc	cccgacaccc	gccaacaccc	gctgacgcgc	7500
cctgacgggc	ttgtctgctc	ccggcatccg	cttacagaca	agctgtgacc	gtctccggga	7560
gctgcatgtg	tcagaggttt	tcaccgtcat	caccgaaacg	cgcgaggcag	ctgcggtaaa	7620
gctcatcagc	gtggtcgtga	agcgattcac	agatgtctgc	ctgttcatcc	gcgtccagct	7680
cgttgagttt	ctccagaagc	gttaatgtct	ggcttctgat	aaagcgggcc	atgttaaggg	7740
cggttttttc	ctgtttggtc	acttgatgcc	tcggtgtaag	ggggaatttc	tgttcatggg	7800
ggtaatgata	ccgatgaaac	gagagaggat	gctcacgata	cgggttactg	atgatgaaca	7860
tgcccggtta	ctggaacggt	gtgagggtaa	acaactggcg	gtatggatgc	ggcgggacca	7920
gagaaaaatc	actcagggtc	aatgccagcg	cttcgttaat	acagatgtag	gtgttccaca	7980
gggtagccag	cagcatcctg	cgatgcagat	ccggaacata	atggtgcagg	gcgctgactt	8040
ccgcgtttcc	agactttacg	aaacacggaa	accgaagacc	attcatgttg	ttgctcaggt	8100
cgcagacggt	ttgcagcagc	agtcgcttca	cgttcgctcg	cgtatcggtg	attcattctg	8160
ctaaccagta	aggcaacccc	gccagcctag	ccgggtcctc	aacgacagga	gcacgatcat	8220
gcgcacccgt	ggccaggacc	caacgctgcc	cgagatgcgc	cgcgtgcggc	tgctggagat	8280
ggcggacgcg	atggatatgt	tctgccaaag	gttggtttgc	gcattcacag	ttctccgcaa	8340
gaattgattg	gctccaattc	ttggagtggg	gaatccgtta	gcgaggtgcc	gccggccttc	8400
attcaggtcg	aggtggcccg	gctccatgca	ccgcgacgca	acgcggggag	gcagacaagg	8460
tataggggcg	cgcctacaat	ccatgccaac	ccgttccatg	tgctcgccga	ggcggcataa	8520
atcgccgtga	cgatcagcgg	tccagtgatc	gaagttaggc	tggttaagagc	cgcgagcgat	8580
ccttgaagct	gtccctgatg	gtcgtcatct	acctgcctgg	acagcatggc	ctgcaacgcg	8640
ggcatcccga	tgccgcggga	agcgagaaga	atcataatgg	ggaaggccat	ccagcctcgc	8700
gtcg						8704

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
25 July 2002 (25.07.2002)

PCT

(10) International Publication Number
WO 02/057476 A3(51) International Patent Classification⁷: **C12N 9/02**,
C12P 13/02, 7/42, C12N 15/52, 15/03(21) International Application Number: **PCT/US02/01887**

(22) International Filing Date: 19 January 2002 (19.01.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/263,053 19 January 2001 (19.01.2001) US(71) Applicant (for all designated States except US): **BASF AKTIENGESELLSCHAFT** [DE/DE]; 67056 Ludwigshafen (DE).

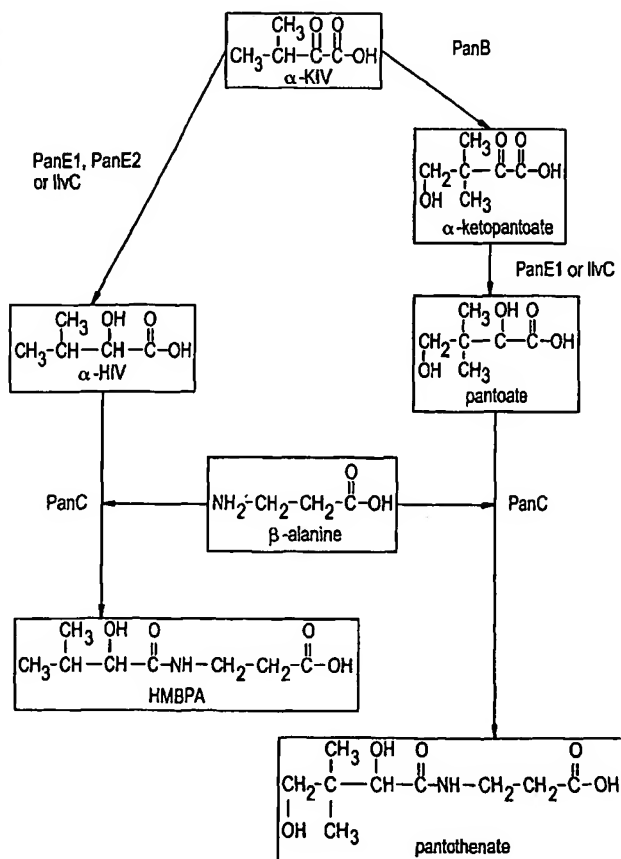
(72) Inventors; and

(75) Inventors/Applicants (for US only): **HERMANN,****Theron** [US/US]; 18 Chilhowie Drive, Kinnelon, NJ 07405 (US). **PATTERSON, Thomas, A.** [US/US]; 89 Church Street, North Attleboro, MA 02760 (US). **PERO, Janice, G.** [US/US]; 20 Solomon Pierce Road, Lexington, MA 02420 (US). **R. ROGERS, Yocum** [US/US]; 4 Orchard Lane, Lexington, MA 02420 (US). **BALDENIUS, Kai-Uwe** [DE/DE]; Kneippstrasse 16, 67063 Ludwigshafen (DE). **BECK, Christine** [DE/DE]; Max-Joseph-Strasse 35, 68167 Mannheim (DE).(74) Agents: **HANLEY, Elizabeth, A.** et al.; Lahive & Cockfield, LLP, 28 State Street, Boston, MA 02109 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

[Continued on next page]

(54) Title: METHODS AND MICROORGANISMS FOR THE PRODUCTION OF 3-(2-HYDROXY-3-METHYL-BUTYRYLAMINO)-PROPIONIC ACID (HMBPA)

(57) Abstract: The present invention features methods of producing 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid ("HMBPA") and α -hydroxyisovalerate (" α -HIV") utilizing microorganisms having modified pantothenate biosynthetic enzyme activities. Recombinant microorganisms and conditions for culturing same are also featured. Also featured are compositions including HMBPA and compositions including α -HIV.



(84) **Designated States (regional):** ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

(88) **Date of publication of the international search report:**
30 October 2003

(15) **Information about Correction:**

Previous Correction:

see PCT Gazette No. 15/2003 of 10 April 2003, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 02/01887

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N9/02 C12P13/02 C12P7/42 C12N15/52 C12N15/03

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12P C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, CHEM ABS Data, WPI Data, PAJ, BIOSIS, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DE 199 07 567 A (KERNFORSCHUNGSANLAGE JUELICH) 24 August 2000 (2000-08-24) page 2, line 29 -page 4, line 24; claims 2,13,15	34-36
E	WO 02 057474 A (BALDENIUS KAI-UWE ;BECK CHRISTINE (DE); PERO JANICE G (US); YOCUM) 25 July 2002 (2002-07-25) the whole document	1-36



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *&* document member of the same patent family

Date of the actual completion of the international search

17 April 2003

Date of mailing of the international search report

02/05/2003

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Piret, B

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 02/01887

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	GORDON E M ET AL: "SYNTHESIS OF SUBSTRATED-BASED INHIBITORS OF HMG COA REDUCTASE (II)" BIOORGANIC & MEDICINAL CHEMISTRY LETTERS, OXFORD, GB, vol. 1, no. 3, 1991, pages 161-164, XP000940927 ISSN: 0960-894X Scheme 1, step b page 162	
A	--- SCHUMMER A ET AL: "POLYFUNCTIONAL (R)-2-HYDROXYCARBOXYLIC ACIDS BY REDUCTION OF 2-OXO ACIDS WITH HYDROGEN GAS OR FORMATE AND RESTING CELLS OF PROTEUS VULGARIS" TETRAHEDRON, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, NL, vol. 47, no. 43, 1991, pages 9019-9034, XP000196414 ISSN: 0040-4020 Compound 19a table 3 -----	21

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 02/01887

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
DE 19907567	A	24-08-2000	DE 19907567 A1	24-08-2000
			CZ 20013037 A3	16-01-2002
			WO 0050624 A1	31-08-2000
			EP 1155139 A1	21-11-2001
			JP 2002537771 A	12-11-2002
			SK 12052001 A3	07-01-2002
<hr/>				
WO 02057474	A	25-07-2002	WO 02061108 A2	08-08-2002
			WO 02057474 A2	25-07-2002
<hr/>				

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
25 July 2002 (25.07.2002)

PCT

(10) International Publication Number
WO 02/057476 A2(51) International Patent Classification⁷: C12P 13/02,
7/42, C12N 15/52, 9/02, 15/03

(21) International Application Number: PCT/US02/01887

(22) International Filing Date: 19 January 2002 (19.01.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/263,053 19 January 2001 (19.01.2001) US(71) Applicant (for all designated States except US): OMNI-
GENE BIOPRODUCTS INC. [US/US]; 763(D) Concord
Avenue, Cambridge, MA 02138 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): HERMANN,
Theron [US/US]; 18 Chilhowie Drive, Kinnelon, NJ
07405 (US). PATTERSON, Thomas, A. [US/US];
89 Church Street, North Attleboro, MA 02760 (US).
PERO, Janice, G. [US/US]; 20 Solomon Pierce Road,
Lexington, MA 02420 (US). R. ROGERS, Yocum
[US/US]; 4 Orchard Lane, Lexington, MA 02420 (US).
BALDENIUS, Kai-Uwe [DE/DE]; Kneippstrasse 16,
67063 Ludwigshafen (DE). BECK, Christine [DE/DE];
Max-Joseph-Strasse 35, 68167 Mannheim (DE).(74) Agents: HANLEY, Elizabeth, A. et al.; Lahive & Cock-
field, LLP, 28 State Street, Boston, MA 02109 (US).(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,
SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ,
VN, YU, ZA, ZM, ZW.(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,
GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,
NE, SN, TD, TG).

Published:

— without international search report and to be republished
upon receipt of that report

(48) Date of publication of this corrected version:

10 April 2003

(15) Information about Correction:

see PCT Gazette No. 15/2003 of 10 April 2003, Section II

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.(54) Title: METHODS AND MICROORGANISMS FOR THE PRODUCTION OF 3-(2-HYDROXY-3-METHYL-BUTYRY-
LAMINO)-PROPIONIC ACID (HMBPA)(57) Abstract: The present invention features methods of producing 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid
("HMBPA") and α -hydroxyisovalerate (" α -HIV") utilizing microorganisms having modified pantothenate biosynthetic enzyme
activities. Recombinant microorganisms and conditions for culturing same are also featured. Also featured are compositions
including HMBPA and compositions including α -HIV.

WO 02/057476 A2

METHODS AND MICROORGANISMS FOR THE PRODUCTION OF 3-(2-HYDROXY-3-METHYL-BUTYRYLAMINO)-PROPIONIC ACID (HMBPA)

Related Applications

5 The present invention claims the benefit of prior-filed provisional Patent Application Serial No. 60/263,053, filed January 19, 2001 (pending). The present invention is also related to U.S. Patent Application Serial No. 09/667,569, filed September 21, 2000 (pending), which is a continuation-in-part of U.S. Patent Application Serial No. 09/400,494, filed September 21, 1999 (abandoned). U.S. Patent Application Serial No. 09/667,569 also claims the benefit of prior-filed provisional Patent Application Serial No. 60/210,072, filed June 7, 2000, provisional Patent Application Serial No. 60/221,836, filed July 28, 2000, and provisional Patent Application Serial No. 60/227,860, filed August 24, 2000. The entire content of each of the above-referenced applications is incorporated herein by this reference.

15

Background of the Invention

Conventional means of synthesizing chemical compounds is *via* synthesis from bulk chemicals, a process which is limited by factors such as substrate availability and/or cost, difficulty in resolving complex mixtures of products, complexities in synthesizing large quantities of compounds in purified form, and difficulty in producing chiral compounds. Accordingly, researchers have recently looked to bacterial or microbial systems that express enzymes useful for various biosynthetic processes, for example, in the synthesis of pharmaceutical compounds, research reagents, nutraceuticals, vitamins, nutritional supplements, antibiotic compounds and the like. In particular, bioconversion processes have been evaluated as a means of favoring production of preferred compounds and recently methods of direct microbial synthesis have been the focus of much research in the areas of pharmaceuticals and agriculture.

25

Summary of the Invention

30 The present invention relates to a processes for the direct microbial synthesis of [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid or 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid ("HMBPA"), referred to interchangeably herein as " β -alanine 2-(*R*)-hydroxyisovalerate", " β -alanine 2-hydroxyisovalerate", " β -alanyl- α -hydroxyisovalerate", N-(2-hydroxy-3-methyl-1-oxobutyl)- β -alanine ("HMOBA")

and/or "pantothenate". In particular, it has been discovered that in microorganisms engineered to overexpress certain enzymes conventionally associated with pantothenate and/or isoleucine-valine (*ilv*) biosynthesis, an alternative biosynthetic pathway is present that competes for key precursors of pantothenate biosynthesis, namely α -ketoisovalerate (5 α -KIV) and β -alanine. α -KIV is converted to α -hydroxyisovalerate (α -HIV) catalyzed by various reductase enzymes and α -HIV is subsequently condensed with β -alanine to produce HMBPA.

In one embodiment, the invention features a process for the production of 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA) that includes culturing 10 a microorganism having increased keto reductase activity or increased pantothenate synthetase activity in the presence of excess α -ketoisovalerate and excess β -alanine, such that HMBPA is produced. In another embodiment, the invention features a process for the production of 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA) that includes culturing a microorganism having increased keto reductase 15 activity and increased pantothenate synthetase activity in the presence of excess α -ketoisovalerate and excess β -alanine, such that HMBPA is produced. In one embodiment, the microorganism has a modified *panE* gene, for example, a modified *panE1* gene and/or a modified *panE2* gene (*e.g.*, the *panE* gene is overexpressed, deregulated or present in multiple copies). In another embodiment, the microorganism 20 has a modified *panC* gene (*e.g.*, the *panC* gene is overexpressed, deregulated or present in multiple copies). In another embodiment, the microorganism further has increased acetohydroxyacid isomeroreductase activity. In another embodiment, the microorganism is cultured under conditions of increased acetohydroxyacid isomeroreductase activity in the presence of excess α -ketoisovalerate and excess β -alanine, such that HMBPA is produced. In yet another embodiment, the microorganism 25 comprises a modified *ilvC* gene (*e.g.*, the *ilvC* gene is overexpressed, deregulated or present in multiple copies). In yet another embodiment, the microorganism further has reduced ketopantoate hydroxymethyltransferase activity (*e.g.*, has a modified *panB* gene, for example a *panB* gene that has been deleted).

30 In another aspect, the invention features a process for the production of 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA) that includes culturing a microorganism having reduced ketopantoate hydroxymethyltransferase activity in the presence of excess α -ketoisovalerate and excess β -alanine, such that HMBPA is produced. In another aspect, the invention features a method for enhancing production

of 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA) relative to pantothenate that includes culturing a recombinant microorganism under conditions such that the HMBPA production is enhanced relative to pantothenate production. In another aspect, the invention features a process for the production of 2-hydroxyisovaleric acid (α-HIV) that includes culturing a microorganism which overexpresses PanE1 or PanE2 and which further has reduced PanC or PanD activity under conditions such that α-HIV is produced. In another aspect, the invention features a process for the production of 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA) that includes culturing a recombinant microorganism having decreased expression or activity of *serA* or *glyA* under conditions such that HMBPA is produced. In another aspect, the invention features a process for the production of 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA) that includes culturing a recombinant microorganism having decreased expression or activity of *serA* and *glyA* under conditions such that HMBPA is produced. Conditions for culturing the above described microorganisms include, for example, conditions of increased steady state glucose, conditions of decreased steady state dissolved oxygen, and/or cultured under conditions of decreased serine. Products produced according to the above described processes and/or methods are also featured. Also featured are recombinant microorganisms utilized in the above-described methods.

Compounds produced according to the methodologies of the present invention have a variety of uses. For example, HMBPA can be used to synthesize inhibitors of HMG CoA Reductase (II) (Gordon *et al. Bio. Med. Chem. Lett.* 1(3):161 (1991). Inhibitors of HMG CoA Reductase (II) have been studied for use as in the treatment of hypercholesterolaemia and coronary atherosclerosis progression. Inhibitors of HMG CoA Reductase also have been used to reduce risk of cardiovascular events in patients at risk. Moreover, the HMBPA precursor 2-hydroxyisovalerate (α-HIV) has been demonstrated to have nutraceutical properties, for example, in the prevention of aging of the skin. In particular, α-hydroxy acids, such as α-HIV (or 2-hydroxyvaline), can be used to synthesize α-hydroxy esters which have been found to induce increased skin thickness by increasing biosyntheses of glycosaminoglycans, proteoglycans, collagen, elastin, and other dermal components. The compounds can be used to treat skin disorders such as age spots, skin lines, wrinkles, photoaging and aging.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

Brief Description of the Drawings

Figure 1 is a schematic representation of the pantothenate and isoleucine-valine (*ilv*) biosynthetic pathways. Pantothenate biosynthetic enzymes are depicted in bold and their corresponding genes indicated in italics. Isoleucine-valine (*ilv*) biosynthetic enzymes are depicted in bold italics and their corresponding genes indicated in italics.

Figure 2 is a schematic representation of the biosynthetic pathway leading to [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid ("HMBPA") in *B. subtilis*.

Figure 3 is a schematic depiction of the structure of [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid ("HMBPA").

Figure 4 is a HPLC chromatogram of a sample of medium from a 14 L fermentation of PA824.

Figure 5 is a mass spectrum depicting the relative monoisotopic mass of [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid.

Figure 6 depicts an alignment of the C-terminal amino acids from known or suspected PanB proteins.

Figure 7 is a schematic representation of the construction of the plasmid pAN624.

Figure 8 is a schematic representation of the construction of the plasmid pAN620.

Figure 9 is a schematic representation of the construction of the plasmid pAN636.

Figure 10 is a schematic representation of the construction of the plasmid pAN637 which allows selection for single or multiple copies using chloramphenicol.

Figure 11 is a schematic representation of the construction of the plasmid pAN238, a plasmid for overexpressing *B. subtilis panE2* from the *P*₂₆ promoter.

Detailed Description of the Invention

The present invention is based, at least in part, on the discovery of a novel biosynthetic pathway in bacteria, namely the [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid ("HMBPA") biosynthetic pathway. In particular, it has been discovered that bacteria are capable of generating HMBPA from α -ketoisovalerate (α -KIV), a key product of the isoleucine-valine (*ilv*) biosynthetic pathway and precursor of the

pantothenate biosynthetic pathway. Production of HMBPA in bacteria involves at least the pantothenate biosynthetic enzymes ketopantoate reductase (the *panE1* gene product) and/or acetohydroxyacid isomeroreductase (the *ilvC* gene product) and results from the condensation of 2-hydroxyisovaleric acid (α -HIV), formed by reduction of α -KIV, and β -alanine, the latter reaction being catalyzed by the pantothenate biosynthetic enzyme pantothenate synthetase (the *panC* gene product). Production of HMBPA is achieved by increasing ketopantoate reductase (*e.g.*, PanE1) and/or PanE2 and/or acetohydroxyacid isomeroreductase activities (*e.g.*, IlvC) in microorganisms, for example, by overexpressing or deregulating the genes encoding said enzymes. Optimal production of HMBPA is achieved by decreasing or deleting ketopantoate hydroxymethyltransferase activity (the *panB* gene product) in microorganisms, for example, by modifying or deleting the *panB* gene which encodes ketopantoate hydroxymethyltransferase (*e.g.*, PanB), optionally in addition to increasing ketopantoate reductase and/or PanE2 and/or acetohydroxyacid isomeroreductase activities in said microorganisms. The substrates α -KIV and β -alanine are required for HMBPA production, the latter provided, for example, by β -alanine feeding and/or increased aspartate- α -decarboxylate activity (the *panD* gene product). Increasing substrate concentration (*i.e.*, α -KIV and/or β -alanine) further enhances production of HMBPA. α -KIV production can be increased by overexpressing *ilvBNCD* genes and/or *alsS*. HMBPA production can further be increased by limiting serine availability or synthesis in appropriately engineered microorganisms.

In order that the present invention may be more readily understood, certain terms are first defined herein.

The term "pantothenate biosynthetic pathway" includes the biosynthetic pathway involving pantothenate biosynthetic enzymes (*e.g.*, polypeptides encoded by biosynthetic enzyme-encoding genes), compounds (*e.g.*, precursors, substrates, intermediates or products), cofactors and the like utilized in the formation or synthesis of pantothenate. The term "pantothenate biosynthetic pathway" includes the biosynthetic pathway leading to the synthesis of pantothenate in microorganisms (*e.g.*, *in vivo*) as well as the biosynthetic pathway leading to the synthesis of pantothenate *in vitro*.

The term "pantothenate biosynthetic enzyme" includes any enzyme utilized in the formation of a compound (*e.g.*, intermediate or product) of the pantothenate biosynthetic pathway. For example, synthesis of pantoate from α -ketoisovalerate (α -KIV) proceeds *via* the intermediate, ketopantoate. Formation of ketopantoate is catalyzed by the pantothenate biosynthetic enzyme ketopantoate hydroxymethyltransferase (the *panB* gene product). Formation of pantoate is catalyzed by the pantothenate biosynthetic enzyme ketopantoate reductase (the *panE* gene product). Synthesis of β -alanine from aspartate is catalyzed by the pantothenate biosynthetic enzyme aspartate- α -decarboxylase (the *panD* gene product). Formation of pantothenate from pantoate and β -alanine (*e.g.*, condensation) is catalyzed by the pantothenate biosynthetic enzyme pantothenate synthetase (the *panC* gene product). Based on the newly discovered HMBPA biosynthesis pathway, pantothenate biosynthetic enzymes may also perform an alternative function as enzymes in the HMBPA biosynthetic pathway described herein.

The term "pantothenate" includes the free acid form of pantothenate, also referred to as "pantothenic acid" as well as any salt thereof (*e.g.*, derived by replacing the acidic hydrogen of pantothenate or pantothenic acid with a cation, for example, calcium, sodium, potassium, ammonium), also referred to as a "pantothenate salt". The term "pantothenate" also includes alcohol derivatives of pantothenate. Preferred pantothenate salts are calcium pantothenate or sodium pantothenate. A preferred alcohol derivative is pantothenol. Pantothenate salts and/or alcohols of the present invention include salts and/or alcohols prepared *via* conventional methods from the free acids described herein. In another embodiment, calcium pantothenate is synthesized directly by a microorganism of the present invention. A pantothenate salt of the present invention can likewise be converted to a free acid form of pantothenate or pantothenic acid by conventional methodology.

The term "isoleucine-valine biosynthetic pathway" includes the biosynthetic pathway involving isoleucine-valine biosynthetic enzymes (*e.g.*, polypeptides encoded by biosynthetic enzyme-encoding genes), compounds (*e.g.*, precursors, substrates, intermediates or products), cofactors and the like utilized in the formation or synthesis of conversion of pyruvate to valine or isoleucine. The term "isoleucine-valine biosynthetic pathway" includes the biosynthetic pathway leading to the synthesis of valine or isoleucine in microorganisms (*e.g.*, *in vivo*) as well as the biosynthetic pathway leading to the synthesis of valine or isoleucine *in vitro*. Figure 1

includes a schematic representation of the isoleucine-valine biosynthetic pathway. Isoleucine-valine biosynthetic enzymes are depicted in bold italics and their corresponding genes indicated in italics

The term “isoleucine-valine biosynthetic enzyme” includes any enzyme
5 utilized in the formation of a compound (*e.g.*, intermediate or product) of the isoleucine-valine biosynthetic pathway. According to Figure 1, synthesis of valine from pyruvate proceeds *via* the intermediates, acetolactate, α,β -dihydroxyisovalerate (α,β -DHIV) and α -ketoisovalerate (α -KIV). Formation of acetolactate from pyruvate is catalyzed by the
10 isoleucine-valine biosynthetic enzyme acetohydroxyacid synthetase (the *ilvBN* gene product, or alternatively, the *alsS* gene product). Formation of α,β -DHIV from acetolactate is catalyzed by the isoleucine-valine biosynthetic enzyme acetohydroxyacidisomero reductase (the *ilvC* gene product). Synthesis of α -KIV from α,β -DHIV is catalyzed by the isoleucine-valine biosynthetic enzyme dihydroxyacid
15 dehydratase (the *ilvD* gene product). Moreover, valine and isoleucine can be interconverted with their respective α -keto compounds by branched chain amino acid transaminases. Based on the newly discovered HMBPA biosynthesis pathway, isoleucine-valine biosynthetic enzymes may also perform an alternative function as
enzymes in the HMBPA biosynthetic pathway described herein.

The term “3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid
20 (“HMBPA”) biosynthetic pathway” includes the alternative biosynthetic pathway involving biosynthetic enzymes and compounds (*e.g.*, substrates and the like) traditionally associated with the pantothenate biosynthetic pathway utilized in the formation or synthesis of HMBPA. The term “HMBPA biosynthetic pathway” includes the biosynthetic pathway leading to the synthesis of HMBPA in microorganisms (*e.g.*, *in vivo*) as well as the biosynthetic pathway leading to the synthesis of HMBPA *in vitro*.
25

The term “HMBPA biosynthetic enzyme” includes any enzyme utilized in the formation of a compound (*e.g.*, intermediate or product) of the HMBPA biosynthetic pathway. For example, synthesis of 2-hydroxyisovaleric acid (α -HIV) from α -ketoisovalerate (α -KIV) is catalyzed by the *panE1* or *panE2* gene product (PanE1,
30 alternatively referred to herein ketopantoate reductase or PanE2, a α -ketoacid reductase that does not significantly contribute to pantothenate biosynthesis) and/or is catalyzed by the *ilvC* gene product (alternatively referred to herein as acetohydroxyacid isomeroreductase). Formation of HMBPA from β -alanine and α -HIV is catalyzed by the *panC* gene product (alternatively referred to herein as pantothenate synthetase).

The term “3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (“HMBPA”)” includes the free acid form of HMBPA, also referred to as “3-(2-hydroxy-3-methyl-butyrylamino)-propionate” as well as any salt thereof (*e.g.*, derived by replacing the acidic hydrogen of [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid or [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionate with a cation, for example, calcium, sodium, potassium, ammonium), also referred to as a “3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid salt” or “HMBPA salt”. Preferred HMBPA salts are calcium HMBPA or sodium HMBPA. HMBPA salts of the present invention include salts prepared *via* conventional methods from the free acids described herein. An HMBPA salt of the present invention can likewise be converted to a free acid form of [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid or [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionate by conventional methodology.

Various aspects of the invention are described in further detail in the following subsections.

I. Targeting Genes Encoding Various Pantothenate and/or Isoleucine-Valine(*ilv*) and/or HMBPA Biosynthetic Enzymes

In one embodiment, the present invention features targeting or modifying various biosynthetic enzymes of the pantothenate and/or isoleucine-valine(*ilv*) and/or HMBPA biosynthetic pathways. In particular, the invention features modifying various enzymatic activities associated with said pathways by modifying or altering the genes encoding said biosynthetic enzymes.

The term “gene”, as used herein, includes a nucleic acid molecule (*e.g.*, a DNA molecule or segment thereof) that, in an organism, can be separated from another gene or other genes, by intergenic DNA (*i.e.*, intervening or spacer DNA which naturally flanks the gene and/or separates genes in the chromosomal DNA of the organism). Alternatively, a gene may slightly overlap another gene (*e.g.*, the 3’ end of a first gene overlapping the 5’ end of a second gene), said overlapping genes separated from other genes by intergenic DNA. A gene may direct synthesis of an enzyme or other protein molecule (*e.g.*, may comprise coding sequences, for example, a contiguous open reading frame (ORF) which encodes a protein) or may itself be functional in the organism. A gene in an organism, may be clustered in an operon, as defined herein, said operon being separated from other genes and/or operons by the intergenic DNA. An “isolated gene”,

as used herein, includes a gene which is essentially free of sequences which naturally flank the gene in the chromosomal DNA of the organism from which the gene is derived (*i.e.*, is free of adjacent coding sequences which encode a second or distinct protein, adjacent structural sequences or the like) and optionally includes 5' and 3' regulatory sequences, for example promoter sequences and/or terminator sequences. In one embodiment, an isolated gene includes predominantly coding sequences for a protein (*e.g.*, sequences which encode *Bacillus* proteins). In another embodiment, an isolated gene includes coding sequences for a protein (*e.g.*, for a *Bacillus* protein) and adjacent 5' and/or 3' regulatory sequences from the chromosomal DNA of the organism from which the gene is derived (*e.g.*, adjacent 5' and/or 3' *Bacillus* regulatory sequences). Preferably, an isolated gene contains less than about 10 kb, 5 kb, 2 kb, 1 kb, 0.5 kb, 0.2 kb, 0.1 kb, 50 bp, 25 bp or 10 bp of nucleotide sequences that naturally flank the gene in the chromosomal DNA of the organism from which the gene is derived.

The term "operon" includes at least two adjacent genes or ORFs, optionally overlapping in sequence at either the 5' or 3' end of at least one gene or ORF. The term "operon" includes a coordinated unit of gene expression that contains a promoter and possibly a regulatory element associated with one or more adjacent genes or ORFs (*e.g.*, structural genes encoding enzymes, for example, biosynthetic enzymes). Expression of the genes (*e.g.*, structural genes) can be coordinately regulated, for example, by regulatory proteins binding to the regulatory element or by anti-termination of transcription. The genes of an operon (*e.g.*, structural genes) can be transcribed to give a single mRNA that encodes all of the proteins.

A "gene having a mutation" or "mutant gene" as used herein, includes a gene having a nucleotide sequence which includes at least one alteration (*e.g.*, substitution, insertion, deletion) such that the polypeptide or protein encoded by said mutant exhibits an activity that differs from the polypeptide or protein encoded by the wild-type nucleic acid molecule or gene. In one embodiment, a gene having a mutation or mutant gene encodes a polypeptide or protein having an increased activity as compared to the polypeptide or protein encoded by the wild-type gene, for example, when assayed under similar conditions (*e.g.*, assayed in microorganisms cultured at the same temperature). As used herein, an "increased activity" or "increased enzymatic activity" is one that is at least 5% greater than that of the polypeptide or protein encoded by the wild-type nucleic acid molecule or gene, preferably at least 5-10% greater, more preferably at least 10-25% greater and even more preferably at least 25-50%, 50-75% or

75-100% greater than that of the polypeptide or protein encoded by the wild-type nucleic acid molecule or gene. Ranges intermediate to the above-recited values, *e.g.*, 75-85%, 85-90%, 90-95%, are also intended to be encompassed by the present invention. As used herein, an “increased activity” or “increased enzymatic activity” can also include an activity that is at least 1.25-fold greater than the activity of the polypeptide or protein encoded by the wild-type gene, preferably at least 1.5-fold greater, more preferably at least 2-fold greater and even more preferably at least 3-fold, 4-fold, 5-fold, 10-fold, 20-fold, 50-fold, 100-fold or greater than the activity of the polypeptide or protein encoded by the wild-type gene.

In another embodiment, a gene having a mutation or mutant gene encodes a polypeptide or protein having a reduced activity as compared to the polypeptide or protein encoded by the wild-type gene, for example, when assayed under similar conditions (*e.g.*, assayed in microorganisms cultured at the same temperature). A mutant gene also can encode no polypeptide or have a reduced level of production of the wild-type polypeptide. As used herein, a “reduced activity” or “reduced enzymatic activity” is one that is at least 5% less than that of the polypeptide or protein encoded by the wild-type nucleic acid molecule or gene, preferably at least 5-10% less, more preferably at least 10-25% less and even more preferably at least 25-50%, 50-75% or 75-100% less than that of the polypeptide or protein encoded by the wild-type nucleic acid molecule or gene. Ranges intermediate to the above-recited values, *e.g.*, 75-85%, 85-90%, 90-95%, are also intended to be encompassed by the present invention. As used herein, a “reduced activity” or “reduced enzymatic activity” can also include an activity that has been deleted or “knocked out” (*e.g.*, approximately 100% less activity than that of the polypeptide or protein encoded by the wild-type nucleic acid molecule or gene).

Activity can be determined according to any well accepted assay for measuring activity of a particular protein of interest. Activity can be measured or assayed directly, for example, measuring an activity of a protein isolated or purified from a cell or microorganism. Alternatively, an activity can be measured or assayed within a cell or microorganism or in an extracellular medium. For example, assaying for a mutant gene (*i.e.*, said mutant encoding a reduced enzymatic activity) can be accomplished by expressing the mutated gene in a microorganism, for example, a mutant microorganism in which the enzyme is temperature-sensitive, and assaying the mutant gene for the ability to complement a temperature sensitive (Ts) mutant for enzymatic activity. A mutant gene that encodes an “increased enzymatic activity” can be one that

complements the Ts mutant more effectively than, for example, a corresponding wild-type gene. A mutant gene that encodes a "reduced enzymatic activity" is one that complements the Ts mutant less effectively than, for example, a corresponding wild-type gene.

5 It will be appreciated by the skilled artisan that even a single substitution in a nucleic acid or gene sequence (*e.g.*, a base substitution that encodes an amino acid change in the corresponding amino acid sequence) can dramatically affect the activity of an encoded polypeptide or protein as compared to the corresponding wild-type polypeptide or protein. A mutant gene (*e.g.*, encoding a mutant polypeptide or protein),
10 as defined herein, is readily distinguishable from a nucleic acid or gene encoding a protein homologue in that a mutant gene encodes a protein or polypeptide having an altered activity, optionally observable as a different or distinct phenotype in a microorganism expressing said mutant gene or producing said mutant protein or polypeptide (*i.e.*, a mutant microorganism) as compared to a corresponding
15 microorganism expressing the wild-type gene. By contrast, a protein homologue has an identical or substantially similar activity, optionally phenotypically indiscernable when produced in a microorganism, as compared to a corresponding microorganism expressing the wild-type gene. Accordingly it is not, for example, the degree of sequence identity between nucleic acid molecules, genes, protein or polypeptides that
20 serves to distinguish between homologues and mutants, rather it is the activity of the encoded protein or polypeptide that distinguishes between homologues and mutants: homologues having, for example, low (*e.g.*, 30-50% sequence identity) sequence identity yet having substantially equivalent functional activities, and mutants, for example sharing 99% sequence identity yet having dramatically different or altered functional
25 activities.

 It will also be appreciated by the skilled artisan that nucleic acid molecules, genes, protein or polypeptides for use in the instant invention can be derived from any microorganisms having a HMBPA biosynthetic pathway, an *ilv* biosynthetic pathway or a pantothenate biosynthetic pathway. Such nucleic acid molecules, genes,
30 protein or polypeptides can be identified by the skilled artisan using known techniques such as homology screening, sequence comparison and the like, and can be modified by the skilled artisan in such a way that expression or production of these nucleic acid molecules, genes, protein or polypeptides occurs in a recombinant microorganism (*e.g.*, by using appropriate promoters, ribosomal binding sites, expression or integration

vectors, modifying the sequence of the genes such that the transcription is increased (taking into account the preferable codon usage), etc., according to techniques described herein and those known in the art).

In one embodiment, the genes of the present invention are derived from a
5 Gram positive microorganism organism (e.g., a microorganism which retains basic dye, for example, crystal violet, due to the presence of a Gram-positive wall surrounding the microorganism). The term "derived from" (e.g., "derived from" a Gram positive microorganism) refers to a gene which is naturally found in the microorganism (e.g., is naturally found in a Gram positive microorganism). In a preferred embodiment, the
10 genes of the present invention are derived from a microorganism belonging to a genus selected from the group consisting of *Bacillus*, *Corynebacterium* (e.g., *Corynebacterium glutamicum*), *Lactobacillus*, *Lactococci* and *Streptomyces*. In a more preferred embodiment, the genes of the present invention are derived from a microorganism is of the genus *Bacillus*. In another preferred embodiment, the genes of the present invention
15 are derived from a microorganism selected from the group consisting of *Bacillus subtilis*, *Bacillus lentimorbus*, *Bacillus lentus*, *Bacillus firmus*, *Bacillus pantothenicus*, *Bacillus amyloliquefaciens*, *Bacillus cereus*, *Bacillus circulans*, *Bacillus coagulans*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus thuringiensis*, *Bacillus halodurans*, and other Group 1 *Bacillus* species, for example, as characterized
20 by 16S rRNA type. In another preferred embodiment, the gene is derived from *Bacillus brevis* or *Bacillus stearothermophilus*. In another preferred embodiment, the genes of the present invention are derived from a microorganism selected from the group consisting of *Bacillus licheniformis*, *Bacillus amyloliquefaciens*, *Bacillus subtilis*, and *Bacillus pumilus*. In a particularly preferred embodiment, the gene is derived from
25 *Bacillus subtilis* (e.g., is *Bacillus subtilis*-derived). The term "derived from *Bacillus subtilis*" or "*Bacillus subtilis*-derived" includes a gene which is naturally found in the microorganism *Bacillus subtilis*. Included within the scope of the present invention are *Bacillus*-derived genes (e.g., *B. subtilis*-derived genes), for example, *Bacillus* or *B. subtilis* *coaX* genes, *serA* genes, *glyA* genes, *coaA* genes, *pan* genes and/or *ilv* genes.

30 In another embodiment, the genes of the present invention are derived from a Gram negative (excludes basic dye) microorganism. In a preferred embodiment, the genes of the present invention are derived from a microorganism belonging to a genus selected from the group consisting of *Salmonella* (e.g., *Salmonella typhimurium*), *Escherichia*, *Klebsiella*, *Serratia*, and *Proteus*. In a more preferred embodiment, the

genes of the present invention are derived from a microorganism of the genus *Escherichia*. In an even more preferred embodiment, the genes of the present invention are derived from *Escherichia coli*. In another embodiment, the genes of the present invention are derived from *Saccharomyces* (e.g., *Saccharomyces cerevisiae*).

5

II. Recombinant Nucleic Acid Molecules and Vectors

The present invention further features recombinant nucleic acid molecules (e.g., recombinant DNA molecules) that include genes described herein (e.g., isolated genes), preferably *Bacillus* genes, more preferably *Bacillus subtilis* genes, even
10 more preferably *Bacillus subtilis* pantothenate biosynthetic genes and/or isoleucine-valine (*ilv*) biosynthetic genes and/or HMBPA biosynthetic genes. The term “recombinant nucleic acid molecule” includes a nucleic acid molecule (e.g., a DNA molecule) that has been altered, modified or engineered such that it differs in nucleotide sequence from the native or natural nucleic acid molecule from which the recombinant
15 nucleic acid molecule was derived (e.g., by addition, deletion or substitution of one or more nucleotides). Preferably, a recombinant nucleic acid molecule (e.g., a recombinant DNA molecule) includes an isolated gene of the present invention operably linked to regulatory sequences. The phrase “operably linked to regulatory sequence(s)” means that the nucleotide sequence of the gene of interest is linked to the regulatory
20 sequence(s) in a manner which allows for expression (e.g., enhanced, increased, constitutive, basal, attenuated, decreased or repressed expression) of the gene, preferably expression of a gene product encoded by the gene (e.g., when the recombinant nucleic acid molecule is included in a recombinant vector, as defined herein, and is introduced into a microorganism).

25 The term “regulatory sequence” includes nucleic acid sequences which affect (e.g., modulate or regulate) expression of other nucleic acid sequences (i.e., genes). In one embodiment, a regulatory sequence is included in a recombinant nucleic acid molecule in a similar or identical position and/or orientation relative to a particular gene of interest as is observed for the regulatory sequence and gene of interest as it
30 appears in nature, e.g., in a native position and/or orientation. For example, a gene of interest can be included in a recombinant nucleic acid molecule operably linked to a regulatory sequence which accompanies or is adjacent to the gene of interest in the natural organism (e.g., operably linked to “native” regulatory sequences (e.g., to the “native” promoter). Alternatively, a gene of interest can be included in a recombinant

nucleic acid molecule operably linked to a regulatory sequence which accompanies or is adjacent to another (*e.g.*, a different) gene in the natural organism. Alternatively, a gene of interest can be included in a recombinant nucleic acid molecule operably linked to a regulatory sequence from another organism. For example, regulatory sequences from
5 other microbes (*e.g.*, other bacterial regulatory sequences, bacteriophage regulatory sequences and the like) can be operably linked to a particular gene of interest.

In one embodiment, a regulatory sequence is a non-native or non-naturally-occurring sequence (*e.g.*, a sequence which has been modified, mutated, substituted, derivatized, deleted including sequences which are chemically synthesized).
10 Preferred regulatory sequences include promoters, enhancers, termination signals, anti-termination signals and other expression control elements (*e.g.*, sequences to which repressors or inducers bind and/or binding sites for transcriptional and/or translational regulatory proteins, for example, in the transcribed mRNA). Such regulatory sequences are described, for example, in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular*
15 *Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989. Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in a microorganism (*e.g.*, constitutive promoters and strong constitutive promoters), those which direct inducible expression of a nucleotide sequence in a microorganism (*e.g.*, inducible
20 promoters, for example, xylose inducible promoters) and those which attenuate or repress expression of a nucleotide sequence in a microorganism (*e.g.*, attenuation signals or repressor sequences). It is also within the scope of the present invention to regulate expression of a gene of interest by removing or deleting regulatory sequences. For example, sequences involved in the negative regulation of transcription can be removed
25 such that expression of a gene of interest is enhanced.

In one embodiment, a recombinant nucleic acid molecule of the present invention includes a nucleic acid sequence or gene that encode at least one bacterial gene product (*e.g.*, a pantothenate biosynthetic enzyme, an isoleucine-valine biosynthetic enzyme and/or a HMBPA biosynthetic enzyme) operably linked to a promoter or
30 promoter sequence. Preferred promoters of the present invention include *Bacillus* promoters and/or bacteriophage promoters (*e.g.*, bacteriophage which infect *Bacillus*). In one embodiment, a promoter is a *Bacillus* promoter, preferably a strong *Bacillus* promoter (*e.g.*, a promoter associated with a biochemical housekeeping gene in *Bacillus* or a promoter associated with a glycolytic pathway gene in *Bacillus*). In another

embodiment, a promoter is a bacteriophage promoter. In a preferred embodiment, the promoter is from the bacteriophage SP01. In a particularly preferred embodiment, a promoter is selected from the group consisting of *P*₁₅, *P*₂₆ or *P*_{veg}, having for example, the following respective sequences:

- 5 GCTATTGACGACAGCTATGGTTCCTGCTCCACCAACCAAACTGTGCTCAGT
ACCGCCAATATTTCTCCCTTGAGGGGTACAAAGAGGTGTCCTAGAAAGAGAT
CCACGCTGTGTAAAAATTTACAAAAAGGTATTGACTTTCCTACAGGGTGT
GTAATAATTTAATTACAGGCGGGGGCAACCCCGCCTGT(SEQ ID NO:1),
GCCTACCTAGCTTCCAAGAAAGATATCCTAACAGCACAAGAGCGGAAAGAT
10 GTTTTGTCTACATCCAGAACAACCTCTGCTAAAATTCCTGAAAAATTTTGCA
AAAAGTTGTTGACTTTATCTACAAGGTGTGGTATAATAATCTTAACAACAGC
AGGACGC (SEQ ID NO:2), and
GAGGAATCATAGAATTTTGTCAAATAATTTTATTGACAACGTCTTATTAAC
GTTGATATAATTTAAATTTTATTGACAAAAATGGGCTCGTGTGTACAATA
15 AATGTAGTGAGGTGGATGCAATG (SEQ ID NO:3). Additional preferred promoters
include *tef* (the translational elongation factor (TEF) promoter) and *pyc* (the pyruvate
carboxylase (PYC) promoter), which promote high level expression in *Bacillus* (e.g.,
Bacillus subtilis). Additional preferred promoters, for example, for use in Gram positive
microorganisms include, but are not limited to, *amy* and SPO2 promoters. Additional
20 preferred promoters, for example, for use in Gram negative microorganisms include, but
are not limited to, *cos*, *tac*, *trp*, *tet*, *trp-tet*, *lpp*, *lac*, *lpp-lac*, *lacIQ*, T7, T5, T3, *gal*, *trc*,
ara, SP6, λ -PR or λ -PL.

- In another embodiment, a recombinant nucleic acid molecule of the
present invention includes a terminator sequence or terminator sequences (e.g.,
25 transcription terminator sequences). The term "terminator sequences" includes
regulatory sequences that serve to terminate transcription of mRNA. Terminator
sequences (or tandem transcription terminators) can further serve to stabilize mRNA
(e.g., by adding structure to mRNA), for example, against nucleases.

- In yet another embodiment, a recombinant nucleic acid molecule of the
30 present invention includes sequences which allow for detection of the vector containing
said sequences (i.e., detectable and/or selectable markers), for example, genes that
encode antibiotic resistance or sequences that overcome auxotrophic mutations, for
example, *trpC*, fluorescent markers, drug markers, and/or colorimetric markers (e.g.,
lacZ/ β -galactosidase). In yet another embodiment, a recombinant nucleic acid molecule

of the present invention includes an artificial ribosome binding site (RBS) or a sequence that becomes transcribed into an artificial RBS. The term "artificial ribosome binding site (RBS)" includes a site within an mRNA molecule (*e.g.*, coded within DNA) to which a ribosome binds (*e.g.*, to initiate translation) which differs from a native RBS (*e.g.*, a RBS found in a naturally-occurring gene) by at least one nucleotide. Preferred artificial RBSs include about 5-6, 7-8, 9-10, 11-12, 13-14, 15-16, 17-18, 19-20, 21-22, 23-24, 25-26, 27-28, 29-30 or more nucleotides of which about 1-2, 3-4, 5-6, 7-8, 9-10, 11-12, 13-15 or more differ from the native RBS (*e.g.*, the native RBS of a gene of interest, for example, the native *panB* RBS TAAACATGAGGAGGAGAAAACATG (SEQ ID NO:4) or the native *panD* RBS ATTCGAGAAATGGAGAGAATATAATATG (SEQ ID NO:5)).

Preferably, nucleotides that differ are substituted such that they are identical to one or more nucleotides of an ideal RBS when optimally aligned for comparisons. Ideal RBSs include, but are not limited to, AGAAAGGAGGTGA (SEQ ID NO:6), TTAAGAAAGGAGGTGANNNNATG (SEQ ID NO:7), TTAGAAAGGAGGTGANNNNNNATG (SEQ ID NO:8), AGAAAGGAGGTGANNNNNNNNATG (SEQ ID NO:9), and AGAAAGGAGGTGANNNNNNNNATG (SEQ ID NO:10). Artificial RBSs can be used to replace the naturally-occurring or native RBSs associated with a particular gene. Artificial RBSs preferably increase translation of a particular gene. Preferred artificial RBSs (*e.g.*, RBSs for increasing the translation of *panB*, for example, of *B. subtilis panB*) include CCCTCTAGAAGGAGGAGAGAAAACATG (SEQ ID NO:11) and CCCTCTAGAGGAGGAGAGAAAACATG (SEQ ID NO:12). Preferred artificial RBSs (*e.g.*, RBSs for increasing the translation of *panD*, for example, of *B. subtilis panD*) include TTAGAAAGGAGGATTAAATATG (SEQ ID NO:13), TTAGAAAGGAGGTTTAATTAATG (SEQ ID NO:14), TTAGAAAGGAGGTGATTAAATG (SEQ ID NO:15), TTAGAAAGGAGGTGTTTAAAATG (SEQ ID NO:16), ATTCGAGAAAGGAGGTGAATATAATATG (SEQ ID NO:17), ATTCGAGAAAGGAGGTGAATAATAATG (SEQ ID NO:18), and ATTCGTAGAAAGGAGGTGAATTAATATG (SEQ ID NO:19).

The present invention further features vectors (*e.g.*, recombinant vectors) that include nucleic acid molecules (*e.g.*, genes or recombinant nucleic acid molecules comprising said genes) as described herein. The term "recombinant vector" includes a vector (*e.g.*, plasmid, phage, phasmid, virus, cosmid or other purified nucleic acid

vector) that has been altered, modified or engineered such that it contains greater, fewer or different nucleic acid sequences than those included in the native or natural nucleic acid molecule from which the recombinant vector was derived. Preferably, the recombinant vector includes a biosynthetic enzyme-encoding gene or recombinant -
5 nucleic acid molecule including said gene, operably linked to regulatory sequences, for example, promoter sequences, terminator sequences and/or artificial ribosome binding sites (RBSs), as defined herein. In another embodiment, a recombinant vector of the present invention includes sequences that enhance replication in bacteria (*e.g.*, replication-enhancing sequences). In one embodiment, replication-enhancing sequences
10 are derived from *E. coli*. In another embodiment, replication-enhancing sequences are derived from pBR322.

In yet another embodiment, a recombinant vector of the present invention includes antibiotic resistance sequences. The term "antibiotic resistance sequences" includes sequences which promote or confer resistance to antibiotics on the host
15 organism (*e.g.*, *Bacillus*). In one embodiment, the antibiotic resistance sequences are selected from the group consisting of *cat* (chloramphenicol resistance) sequences, *tet* (tetracycline resistance) sequences, *erm* (erythromycin resistance) sequences, *neo* (neomycin resistance) sequences, *kan* (kanamycin resistance) and *spec* (spectinomycin resistance) sequences. Recombinant vectors of the present invention can further include
20 homologous recombination sequences (*e.g.*, sequences designed to allow recombination of the gene of interest into the chromosome of the host organism). For example, *bpr*, *vpr*, and/or *amyE* sequences can be used as homology targets for recombination into the host chromosome. It will further be appreciated by one of skill in the art that the design of a vector can be tailored depending on such factors as the choice of microorganism to
25 be genetically engineered, the level of expression of gene product desired and the like.

IV. Recombinant Microorganisms

The present invention further features microorganisms, *i.e.*, recombinant
30 microorganisms, that include vectors or genes (*e.g.*, wild-type and/or mutated genes) as described herein. As used herein, the term "recombinant microorganism" includes a microorganism (*e.g.*, bacteria, yeast cell, fungal cell, etc.) that has been genetically altered, modified or engineered (*e.g.*, genetically engineered) such that it exhibits an altered, modified or different genotype and/or phenotype (*e.g.*, when the genetic

modification affects coding nucleic acid sequences of the microorganism) as compared to the naturally-occurring microorganism from which it was derived.

In one embodiment, a recombinant microorganism of the present invention is a Gram positive organism (e.g., a microorganism which retains basic dye, for example, crystal violet, due to the presence of a Gram-positive wall surrounding the microorganism). In a preferred embodiment, the recombinant microorganism is a microorganism belonging to a genus selected from the group consisting of *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Lactococci* and *Streptomyces*. In a more preferred embodiment, the recombinant microorganism is of the genus *Bacillus*. In another preferred embodiment, the recombinant microorganism is selected from the group consisting of *Bacillus subtilis*, *Bacillus lentimorbus*, *Bacillus lentus*, *Bacillus firmus*, *Bacillus pantothenicus*, *Bacillus amyloliquefaciens*, *Bacillus cereus*, *Bacillus circulans*, *Bacillus coagulans*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus thuringiensis*, *Bacillus halodurans*, and other Group 1 *Bacillus* species, for example, as characterized by 16S rRNA type. In another preferred embodiment, the recombinant microorganism is *Bacillus brevis* or *Bacillus stearothermophilus*. In another preferred embodiment, the recombinant microorganism is selected from the group consisting of *Bacillus licheniformis*, *Bacillus amyloliquefaciens*, *Bacillus subtilis*, and *Bacillus pumilus*.

In another embodiment, the recombinant microorganism is a Gram negative (excludes basic dye) organism. In a preferred embodiment, the recombinant microorganism is a microorganism belonging to a genus selected from the group consisting of *Salmonella*, *Escherichia*, *Klebsiella*, *Serratia*, and *Proteus*. In a more preferred embodiment, the recombinant microorganism is of the genus *Escherichia*. In an even more preferred embodiment, the recombinant microorganism is *Escherichia coli*. In another embodiment, the recombinant microorganism is *Saccharomyces* (e.g., *S. cerevisiae*).

A preferred "recombinant" microorganism of the present invention is a microorganism having a deregulated pantothenate biosynthesis pathway or enzyme, a deregulated isoleucine-valine (*ilv*) biosynthetic pathway or enzyme and/or a deregulated HMBPA biosynthetic pathway or enzyme. The term "deregulated" or "deregulation" includes the alteration or modification of at least one gene in a microorganism that encodes an enzyme in a biosynthetic pathway, such that the level or activity of the biosynthetic enzyme in the microorganism is altered or modified. Preferably, at least

one gene that encodes an enzyme in a biosynthetic pathway is altered or modified such that the gene product is enhanced or increased. The phrase "deregulated pathway" can also include a biosynthetic pathway in which more than one gene that encodes an enzyme in a biosynthetic pathway is altered or modified such that the level or activity of more than one biosynthetic enzyme is altered or modified. The ability to "deregulate" a pathway (*e.g.*, to simultaneously deregulate more than one gene in a given biosynthetic pathway) in a microorganism in some cases arises from the particular phenomenon of microorganisms in which more than one enzyme (*e.g.*, two or three biosynthetic enzymes) are encoded by genes occurring adjacent to one another on a contiguous piece of genetic material termed an "operon" (defined herein). Due to the coordinated regulation of genes included in an operon, alteration or modification of the single promoter and/or regulatory element can result in alteration or modification of the expression of each gene product encoded by the operon. Alteration or modification of the regulatory element can include, but is not limited to removing the endogenous promoter and/or regulatory element(s), adding strong promoters, inducible promoters or multiple promoters or removing regulatory sequences such that expression of the gene products is modified, modifying the chromosomal location of the operon, altering nucleic acid sequences adjacent to the operon or within the operon such as a ribosome binding site, increasing the copy number of the operon, modifying proteins (*e.g.*, regulatory proteins, suppressors, enhancers, transcriptional activators and the like) involved in transcription of the operon and/or translation of the gene products of the operon, or any other conventional means of deregulating expression of genes routine in the art (including but not limited to use of antisense nucleic acid molecules, for example, to block expression of repressor proteins). Deregulation can also involve altering the coding region of one or more genes to yield, for example, an enzyme that is feedback resistant or has a higher or lower specific activity.

In another preferred embodiment, a recombinant microorganism is designed or engineered such that at least one pantothenate biosynthetic enzyme, at least one isoleucine-valine biosynthetic enzyme, and/or at least one HMBPA biosynthetic enzyme is overexpressed. The term "overexpressed" or "overexpression" includes expression of a gene product (*e.g.*, a biosynthetic enzyme) at a level greater than that expressed prior to manipulation of the microorganism or in a comparable microorganism which has not been manipulated. In one embodiment, the microorganism can be

genetically designed or engineered to overexpress a level of gene product greater than that expressed in a comparable microorganism which has not been engineered.

Genetic engineering can include, but is not limited to, altering or modifying regulatory sequences or sites associated with expression of a particular gene (e.g., by adding strong promoters, inducible promoters or multiple promoters or by removing regulatory sequences such that expression is constitutive), modifying the chromosomal location of a particular gene, altering nucleic acid sequences adjacent to a particular gene such as a ribosome binding site, increasing the copy number of a particular gene, modifying proteins (e.g., regulatory proteins, suppressors, enhancers, transcriptional activators and the like) involved in transcription of a particular gene and/or translation of a particular gene product, or any other conventional means of deregulating expression of a particular gene routine in the art (including but not limited to use of antisense nucleic acid molecules, for example, to block expression of repressor proteins). Genetic engineering can also include deletion of a gene, for example, to block a pathway or to remove a repressor. In embodiments featuring microorganisms having deleted genes, the skilled artisan will appreciate that at least low levels of certain compounds may be required to be present in or added to the culture medium in order that the viability of the microorganism is not compromised. Often, such low levels are present in complex culture media as routinely formulated. Moreover, in processes featuring culturing microorganisms having deleted genes cultured under conditions such that commercially or industrially attractive quantities of product are produced, it may be necessary to supplement culture media with slightly increased levels of certain compounds. For example, in processes featuring culturing a microorganism having a deleted *panB* gene, at least low levels of pantothenate must be present in the media, e.g., levels such as those found in routinely formulated complex media, whereas slightly increased levels of pantothenate may be added to the media in order to produce commercially or industrially attractive amounts of, for example, HMBPA. For example, 10-30 mg/L pantothenate can be added to the media in order to produce commercially or industrially attractive amounts of HMBPA.

In another embodiment, the microorganism can be physically or environmentally manipulated to overexpress a level of gene product greater than that expressed prior to manipulation of the microorganism or in a comparable microorganism which has not been manipulated. For example, a microorganism can be treated with or cultured in the presence of an agent known or suspected to increase transcription of a

particular gene and/or translation of a particular gene product such that transcription and/or translation are enhanced or increased. Alternatively, a microorganism can be cultured at a temperature selected to increase transcription of a particular gene and/or translation of a particular gene product such that transcription and/or translation are enhanced or increased.

V. Culturing and Fermenting Recombinant Microorganisms

The term "culturing" includes maintaining and/or growing a living microorganism of the present invention (*e.g.*, maintaining and/or growing a culture or strain). In one embodiment, a microorganism of the invention is cultured in liquid media. In another embodiment, a microorganism of the invention is cultured in solid media or semi-solid media. In a preferred embodiment, a microorganism of the invention is cultured in media (*e.g.*, a sterile, liquid media) comprising nutrients essential or beneficial to the maintenance and/or growth of the microorganism (*e.g.*, carbon sources or carbon substrate, for example carbohydrate, hydrocarbons, oils, fats, fatty acids, organic acids, and alcohols; nitrogen sources, for example, peptone, yeast extracts, meat extracts, malt extracts, urea, ammonium sulfate, ammonium chloride, ammonium nitrate and ammonium phosphate; phosphorus sources, for example, phosphoric acid, sodium and potassium salts thereof; trace elements, for example, magnesium, iron, manganese, calcium, copper, zinc, boron, molybdenum, and/or cobalt salts; as well as growth factors such as amino acids, vitamins, growth promoters and the like).

Preferably, microorganisms of the present invention are cultured under controlled pH. The term "controlled pH" includes any pH which results in production of the desired product (*e.g.*, HMBPA). In one embodiment microorganisms are cultured at a pH of about 7. In another embodiment, microorganisms are cultured at a pH of between 6.0 and 8.5. The desired pH may be maintained by any number of methods known to those skilled in the art.

Also preferably, microorganisms of the present invention are cultured under controlled aeration. The term "controlled aeration" includes sufficient aeration (*e.g.*, oxygen) to result in production of the desired product (*e.g.*, HMBPA). In one embodiment, aeration is controlled by regulating oxygen levels in the culture, for example, by regulating the amount of oxygen dissolved in culture media. Preferably, aeration of the culture is controlled by agitating the culture. Agitation may be provided by a propeller or similar mechanical agitation equipment, by revolving or shaking the

culture vessel (*e.g.*, tube or flask) or by various pumping equipment. Aeration may be further controlled by the passage of sterile air or oxygen through the medium (*e.g.*, through the fermentation mixture). Also preferably, microorganisms of the present invention are cultured without excess foaming (*e.g.*, *via* addition of antifoaming agents).

- 5 Moreover, microorganisms of the present invention can be cultured under controlled temperatures. The term "controlled temperature" includes any temperature which results in production of the desired product (*e.g.*, HMBPA). In one embodiment, controlled temperatures include temperatures between 15°C and 95°C. In another embodiment, controlled temperatures include temperatures between 15°C and 70°C.
- 10 Preferred temperatures are between 20°C and 55°C, more preferably between 30°C and 50°C.

- Microorganisms can be cultured (*e.g.*, maintained and/or grown) in liquid media and preferably are cultured, either continuously or intermittently, by conventional culturing methods such as standing culture, test tube culture, shaking culture (*e.g.*, rotary shaking culture, shake flask culture, etc.), aeration spinner culture, or fermentation.
- 15 In a preferred embodiment, the microorganisms are cultured in shake flasks. In a more preferred embodiment, the microorganisms are cultured in a fermentor (*e.g.*, a fermentation process). Fermentation processes of the present invention include, but are not limited to, batch, fed-batch and continuous processes or methods of fermentation.
- 20 The phrase "batch process" or "batch fermentation" refers to a closed system in which the composition of media, nutrients, supplemental additives and the like is set at the beginning of the fermentation and not subject to alteration during the fermentation, however, attempts may be made to control such factors as pH and oxygen concentration to prevent excess media acidification and/or microorganism death. The phrase "fed-
- 25 batch process" or "fed-batch" fermentation refers to a batch fermentation with the exception that one or more substrates or supplements are added (*e.g.*, added in increments or continuously) as the fermentation progresses. The phrase "continuous process" or "continuous fermentation" refers to a system in which a defined fermentation media is added continuously to a fermentor and an equal amount of used or
- 30 "conditioned" media is simultaneously removed, preferably for recovery of the desired product (*e.g.*, HMBPA). A variety of such processes have been developed and are well-known in the art.

The phrase "culturing under conditions such that a desired compound is produced" includes maintaining and/or growing microorganisms under conditions (*e.g.*, temperature, pressure, pH, duration, etc.) appropriate or sufficient to obtain production of the desired compound or to obtain desired yields of the particular compound being produced. For example, culturing is continued for a time sufficient to produce the desired amount of a compound (*e.g.*, HMBPA). Preferably, culturing is continued for a time sufficient to substantially reach suitable production of the compound (*e.g.*, a time sufficient to reach a suitable concentration of HMBPA or suitable ratio of HMBPA:pantothenate). In one embodiment, culturing is continued for about 12 to 24 hours. In another embodiment, culturing is continued for about 24 to 36 hours, 36 to 48 hours, 48 to 72 hours, 72 to 96 hours, 96 to 120 hours, 120 to 144 hours, or greater than 144 hours. In yet another embodiment, microorganisms are cultured under conditions such that at least about 5 to 10 g/L of compound are produced in about 36 hours, at least about 10 to 20 g/L compound are produced in about 48 hours, or at least about 20 to 30 g/L compound in about 72 hours. In yet another embodiment, microorganisms are cultured under conditions such that at least a ratio of HMBPA:HMBPA+pantothenate of 1:10 is achieved (*i.e.*, 10% HMBPA versus 90% pantothenate, for example, as determined by comparing the peaks when a sample of product is analyzed by HPLC), preferably such that at least a ratio of 2:10 is achieved (20% HMBPA versus 90% pantotheante), more preferably such that a ratio of at least 2.5:10 is achieved (25% HMBPA versus 75% pantotheante), more preferably at least 3:10 (30% HMBPA versus 70% pantotheante), 4:10 (40% HMBPA versus 60% pantotheante), 5:10 (50% HMBPA versus 50% pantotheante), 6:10 (60% HMBPA versus 40% pantotheante), 7:10 (70% HMBPA versus 30% pantotheante), 8:10 (80% HMBPA versus 20% pantotheante), 9:10 (90% HMBPA versus 10% pantotheante) or greater.

The methodology of the present invention can further include a step of recovering a desired compound (*e.g.*, HMBPA). The term "recovering" a desired compound includes extracting, harvesting, isolating or purifying the compound from culture media. Recovering the compound can be performed according to any conventional isolation or purification methodology known in the art including, but not limited to, treatment with a conventional resin (*e.g.*, anion or cation exchange resin, non-ionic adsorption resin, etc.), treatment with a conventional adsorbent (*e.g.*, activated charcoal, silicic acid, silica gel, cellulose, alumina, etc.), alteration of pH, solvent extraction (*e.g.*, with a conventional solvent such as an alcohol, ethyl acetate, hexane and

the like), dialysis, filtration, concentration, crystallization, recrystallization, pH adjustment, lyophilization and the like. For example, a compound can be recovered from culture media by first removing the microorganisms from the culture. Media are then passed through or over a cation exchange resin to remove cations and then through
5 or over an anion exchange resin to remove inorganic anions and organic acids having stronger acidities than the compound of interest. The resulting compound can subsequently be converted to a salt (*e.g.*, a calcium salt) as described herein.

Preferably, a desired compound of the present invention is "extracted", "isolated" or "purified" such that the resulting preparation is substantially free of other
10 media components (*e.g.*, free of media components and/or fermentation byproducts). The language "substantially free of other media components" includes preparations of the desired compound in which the compound is separated from media components or fermentation byproducts of the culture from which it is produced. In one embodiment, the preparation has greater than about 80% (by dry weight) of the desired compound
15 (*e.g.*, less than about 20% of other media components or fermentation byproducts), more preferably greater than about 90% of the desired compound (*e.g.*, less than about 10% of other media components or fermentation byproducts), still more preferably greater than about 95% of the desired compound (*e.g.*, less than about 5% of other media components or fermentation byproducts), and most preferably greater than about 98-99% desired
20 compound (*e.g.*, less than about 1-2% other media components or fermentation byproducts). When the desired compound has been derivatized to a salt, the compound is preferably further free of chemical contaminants associated with the formation of the salt. When the desired compound has been derivatized to an alcohol, the compound is
25 preferably further free of chemical contaminants associated with the formation of the alcohol.

In an alternative embodiment, the desired compound is not purified from the microorganism, for example, when the microorganism is biologically non-hazardous (*e.g.*, safe). For example, the entire culture (or culture supernatant) can be used as a source of product (*e.g.*, crude product). In one embodiment, the culture (or culture
30 supernatant) is used without modification. In another embodiment, the culture (or culture supernatant) is concentrated. In yet another embodiment, the culture (or culture supernatant) is dried or lyophilized.

Preferably, a production method of the present invention results in production of the desired compound at a significantly high yield. The phrase “significantly high yield” includes a level of production or yield which is sufficiently elevated or above what is usual for comparable production methods, for example, which is elevated to a level sufficient for commercial production of the desired product (*e.g.*, production of the product at a commercially feasible cost). In one embodiment, the invention features a production method that includes culturing a recombinant microorganism under conditions such that the desired product (*e.g.*, HMBPA) is produced at a level greater than 2 g/L. In another embodiment, the invention features a production method that includes culturing a recombinant microorganism under conditions such that the desired product (*e.g.*, HMBPA) is produced at a level greater than 10 g/L. In another embodiment, the invention features a production method that includes culturing a recombinant microorganism under conditions such that the desired product (*e.g.*, HMBPA) is produced at a level greater than 20 g/L. In yet another embodiment, the invention features a production method that includes culturing a recombinant microorganism under conditions such that the desired product (*e.g.*, HMBPA) is produced at a level greater than 30 g/L. In yet another embodiment, the invention features a production method that includes culturing a recombinant microorganism under conditions such that the desired product (*e.g.*, HMBPA) is produced at a level greater than 40 g/L. The invention further features a production method for producing the desired compound that involves culturing a recombinant microorganism under conditions such that a sufficiently elevated level of compound is produced within a commercially desirable period of time.

Depending on the biosynthetic enzyme or combination of biosynthetic enzymes manipulated, it may be desirable or necessary to provide (*e.g.*, feed) microorganisms of the present invention at least one biosynthetic precursor such that the desired compound or compounds are produced. The term “biosynthetic precursor” or “precursor” includes an agent or compound which, when provided to, brought into contact with, or included in the culture medium of a microorganism, serves to enhance or increase biosynthesis of the desired product. In one embodiment, the biosynthetic precursor or precursor is aspartate. In another embodiment, the biosynthetic precursor or precursor is β -alanine. The amount of aspartate or β -alanine added is preferably an amount that results in a concentration in the culture medium sufficient to enhance productivity of the microorganism (*e.g.*, a concentration sufficient to enhance production

of HMBPA. The term "excess β -alanine" includes β -alanine levels increased or higher than those routinely utilized for culturing the microorganism in question. For example, culturing the *Bacillus* microorganisms described in the instant Examples is routinely done in the presence of about 0-5 g/L β -alanine. Accordingly, excess β -alanine levels can include levels of about 5-10 g/L or more preferably about 5-20 g/L β -alanine. Biosynthetic precursors of the present invention can be added in the form of a concentrated solution or suspension (e.g., in a suitable solvent such as water or buffer) or in the form of a solid (e.g., in the form of a powder). Moreover, biosynthetic precursors of the present invention can be added as a single aliquot, continuously or intermittently over a given period of time.

In yet another embodiment, the biosynthetic precursor is valine. In yet another embodiment, the biosynthetic precursor is α -ketoisovalerate. Preferably, valine or α -ketoisovalerate is added in an amount that results in a concentration in the medium sufficient for production of the desired product (e.g., HMBPA) to occur. The term "excess α -KIV" includes α -KIV levels increased or higher than those routinely utilized for culturing the microorganism in question. For example, culturing the *Bacillus* microorganisms described in the instant Examples can be done in the presence of about 0-5 g/L α -KIV. Accordingly, excess α -KIV levels can include levels of about 5-10 g/L, and more preferably about 5-20 g/L. The term "excess valine" includes valine levels increased or higher than those routinely utilized for culturing the microorganism in question. For example, culturing the *Bacillus* microorganisms described in the instant Examples is routinely done in the presence of about 0-0.5 g/L valine. Accordingly, excess valine levels can include levels of about 0.5-5 g/L, preferably about 5-20 g/L valine. Biosynthetic precursors are also referred to herein as "supplemental biosynthetic substrates".

Moreover, certain aspects of the present invention include culturing microorganisms (e.g., recombinant microorganisms) under conditions of increased steady state glucose, decreased steady state dissolved oxygen and/or decreased serine. The term "increased steady state glucose" includes steady state glucose levels increased or higher than those routinely utilized for culturing the microorganism in question. For example, culturing the *Bacillus* microorganisms described in the instant Examples is routinely done in the presence of about 0.2-1.0 g/L steady state glucose. Accordingly, increased steady state glucose levels can include levels of about 1-2 g/l, about 2-5 g/l, and preferably about 5-20 g/L steady state glucose. The term "decreased steady state

dissolved oxygen” includes steady state dissolved oxygen levels less or lower than those routinely utilized for culturing the microorganism in question and, for example, inversely correlates with increased steady state glucose levels. For example, culturing the *Bacillus* microorganisms described in the instant Examples is routinely done in the presence of about 10-30% dissolved oxygen. Accordingly, decreased steady state dissolved oxygen can include levels of about 0-10%, and preferably about 0-5% steady state dissolved oxygen. The term “reduced serine” includes serine levels within the lower range of those routinely utilized for culturing the microorganism in question. For example, culturing the *Bacillus* microorganisms described in the instant Examples is routinely done in the presence of about 0-0.5 g/L serine. Accordingly, reduced serine levels can include, for example, levels of 0-0.1 g/L serine.

Another aspect of the present invention includes biotransformation processes which feature the recombinant microorganisms described herein. The term “biotransformation process”, also referred to herein as “bioconversion processes”, includes biological processes which results in the production (*e.g.*, transformation or conversion) of appropriate substrates and/or intermediate compounds into a desired product (*e.g.*, HMBPA).

In one embodiment, the invention features a biotransformation process for the production of HMBPA comprising contacting a microorganism which overexpresses a reductase (*e.g.*, overexpresses PanE1, PanE2 and/or IlvC) with appropriate substrates or precursors under conditions such that HMBPA is produced and recovering said HMBPA. In another embodiment, the invention features a biotransformation process for the production of HMBPA comprising contacting a microorganism which has a reduced or deleted PanB activity with appropriate substrates or precursors under conditions such that HMBPA is produced and recovering said HMBPA. In yet another embodiment, the invention features a biotransformation process for the production of HMBPA comprising contacting a microorganism which overexpresses at least one reductase and has a reduced or deleted PanB activity with appropriate substrates or precursors under conditions such that HMBPA is produced and recovering said HMBPA. Preferred recombinant microorganisms for carrying out the above-described biotransformations include the recombinant microorganisms described herein. In yet another embodiment, the invention features a biotransformation reaction that includes contacting α -HIV and β -alanine with isolated or purified PanC under conditions such that HMBPA is produced. α -HIV can optionally be obtained by

contacting α -KIV with purified or isolated reductase (*e.g.*, PanE1, PanE2 and/or IlvC) and a source of reducing equivalent, for example, NADH. Conditions under which α -HIV or HMBPA are produced can include any conditions which result in the desired production of α -HIV or HMBPA, respectively. In yet another embodiment, the present invention includes a method of producing α -HIV that includes culturing a microorganism that overexpresses PanE1 and/or PanE2, and/or IlvC and has a reduced or deleted PanC or PanD (to reduce HMBPA or β -alanine synthesis, respectively) under conditions such that α -HIV is produced.

The microorganism(s) and/or enzymes used in the biotransformation reactions are in a form allowing them to perform their intended function (*e.g.*, producing a desired compound). The microorganisms can be whole cells, or can be only those portions of the cells necessary to obtain the desired end result. The microorganisms can be suspended (*e.g.*, in an appropriate solution such as buffered solutions or media), rinsed (*e.g.*, rinsed free of media from culturing the microorganism), acetone-dried, immobilized (*e.g.*, with polyacrylamide gel or k-carrageenan or on synthetic supports, for example, beads, matrices and the like), fixed, cross-linked or permeablized (*e.g.*, have permeablized membranes and/or walls such that compounds, for example, substrates, intermediates or products can more easily pass through said membrane or wall).

20

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application are incorporated herein by reference.

25

EXAMPLES

Example I: Discovery and Characterization of the [R]-3-(2-hydroxy-3-methyl-butrylamino)-propionic acid (HMBPA) Biosynthetic Pathway

5 In developing *Bacillus* strains for the production of pantothenate, various genetic manipulations were made to enzymes involved in the pantothenate biosynthetic pathway and the isoleucine-valine (*ilv*) pathway (Figure 1) as described in U.S. Patent Application Serial No. 09/400,494 and U.S. Patent Application Serial No. 09/667,569. For example, strains having a deregulated *panBCD* operon and/or having deregulated
10 *panE1* exhibited enhanced pantothenate production (when cultured in the presence of β -alanine and α -KIV). Strains further deregulated for *ilvBNC* and *ilvD* exhibited enhanced pantothenate production in the presence of only β -alanine. Moreover, it was possible to achieve β -alanine independence by further deregulating *panD*.

An exemplary strain is PA824, a tryptophan prototroph, Spec and Tet
15 resistant, deregulated for *panBCD* at the *panBCD* locus, deregulated for *panE1* at the *panE1* locus (two genes in the *B. subtilis* genome are homologous to *E. coli panE*, *panE1* and *panE2*, the former encoding the major ketopantoate reductase involved in pantothenate production, while *panE2* does not contribute to pantothenate synthesis (U.S. Patent Application Serial No. 09/400,494), deregulated for *ilvD* at the *ilvD* locus,
20 overexpressing an *ilvBNC* cassette at the *amyE* locus, and overexpressing *panD* at the *bpr* locus.

The production of pantothenic acid by PA824 was investigated in 14 L fermentor vessels. The composition of the batch and feed media are as follows.

25

BATCH

	MATERIAL	g/L (final)
1	Yeast extract	10
2	Na Glutamate	5
3	(NH ₄) ₂ SO ₄	8
4	KH ₂ PO ₄	5
5	K ₂ HPO ₄	7.6

Added After Sterilization and Cool Down

1	Glucose	2.5
2	CaCl ₂	0.1
3	MgCl ₂	1
4	Sodium Citrate	1
5	FeSO ₄ ·7 H ₂ O	0.01
5	SM-1000X	1 ml

The final volume of the batch medium is 6 L. The trace element solution SM-1000X has following composition: 0.15 g Na₂MoO₄·2 H₂O, 2.5 g H₃BO₃, 0.7 g CoCl₂·6 H₂O, 0.25 g CuSO₄·5 H₂O, 1.6 g MnCl₂·4 H₂O, 0.3 g ZnSO₄·7 H₂O are dissolved in water (final volume 1L).

The batch medium was inoculated with 60 ml of shake flask PA824 culture (OD=10 in SVY medium: Difco Veal Infusion broth 25 g, Difco Yeast extract 5 g, Sodium Glutamate 5 g, (NH₄)₂SO₄ 2.7 g in 740 ml H₂O, autoclave; add 200 ml sterile 1 M K₂HPO₄ (pH 7) and 60 ml sterile 50% Glucose solution (final volume 1L)). The fermentation was run at 43 °C at an air flow rate of 12 L/min as a glucose limited fed batch. The initial batched glucose (2.5 g/L) was consumed during exponential growth. Afterwards glucose concentrations were maintained between 0.2-1 g/L by continuous feeding of FEED solution as follows.

FEED

	MATERIAL	g/L (final)
1	Glucose	550
2	CaCl ₂	0.1
3	SM-1000X	3 ml

15

The variable feed rate pump was computer controlled and linked to the glucose concentration in the tank by an algorithm. In this example the total feeding was 6 L.

During fermentation the pH was set at 7.2. Control was achieved by pH measurements linked to computer control. The pH value was maintained by feeding either a 25% NH₃-solution or a 20% H₃PO₄-solution. NH₃ acts simultaneously as a N-

20

source for the fermentation. The dissolved oxygen concentration $[pO_2]$ was set at 30% by regulation of the agitation and aeration rate. Foaming was controlled by addition of silicone oil. After the stop of the addition of the feed solution, in this example after 48 h, the fermentation was continued until the $[pO_2]$ value reached 95%. Then the fermentation was stopped by killing the microorganism through sterilization for 30 min. The successful sterilization was proven by plating a sample of the fermentation broth on agar plates. The pantothenate titer in the fermentation broth was 21.7 g/L after sterilization and removal of the cells by centrifugation (determined by HPLC analysis).

For HPLC analysis the fermentation broth sample was diluted with sterile water (1:40). 5 μ l of this dilution was injected into a HPLC column (Aqua C18, 5 μ m, 150*2.0 mm, PhenomenexTM). Temperature of the column was held at 40°C. Mobile phase A was 14.8 mM H_3PO_3 , mobile phase B 100% Acetonitrile. Flow rate was constant at 0.5 mL/min. A gradient was applied :

	start:	2% mobile phase B
15	0-3 min	linear increase to 3% mobile phase B
	3-3.5 min	linear increase to 20% mobile phase B

The detection was carried out by an UV-detector (210 nm). Run time was 7 min with an additional 3 min posttime. The retention time for pantothenic acid is 3.9 minutes. The HPLC chromatogram for the above mentioned sample is given in Figure 4.

Identification of compound related to peak with retention time 4.7 minutes

Under the described fermentation conditions, PA824 routinely yields approximately 20-30 g/L pantothenate. In addition to producing significant quantities of pantothenate, it was discovered a second compound eluted with an approximate retention time of 4.7 minutes in this system. The second prominent product formed in the fermentation was shown to be [R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA) (also referred to herein as “ β -alanine 2-(R)-hydroxyisovalerate”, “ β -alanine 2-hydroxyisovalerate”, and/or “ β -alanyl- α -hydroxyisovalerate). It was identified by its mass spectrum (Figure 5; relative monoisotopic mass 189), 1H - and ^{13}C -NMR (data not shown) after chromatographic purification by reverse phase flash chromatography (mobile phase 10 mM KH_2PO_4 , with increasing contents of acetonitrile (1-50%)).

In order to verify the identity of the compound, deliberate synthesis of racemic β -alanine 2-hydroxyisovalerate was performed as follows. β -alanine (2.73 g / 30 mmol) and sodium methoxide (5.67 g of a 30% solution in methanol / 31.5 mmol) were dissolved in methanol (40 mL). Methyl 2-hydroxyisovalerate (2-hydroxy-3-methylbutyric acid methyl ester) (3.96g / 30 mmol) was added and refluxed for 18 hours. Methanol was then removed by rotavap and replaced by tert-butanol (50 mL). Potassium tert-butoxide was added (50 mg) and refluxed for 26 hours. The solvent was removed *in vacuo*, the residue dissolved in water (50 mL) and passed through a strongly acidic ion-exchange resin (H^+ -form Lewatite™ S 100 G1; 100 mL). More water is used to rinse the ion exchanger. The aqueous eluates are combined and the water removed *in vacuo*. The residue is subjected to flash chromatography (silica gel; 2% acetic acid in ethyl acetate as eluent) and the product fractions evaporated to give a solid residue. The residue was recrystallized from ethyl acetate / toluene (10 mL / 20 mL, respectively) and analytically pure HMBPA (β -alanine 2-hydroxyisovalerate) was obtained, which showed a relative monoisotopic mass of 190 ($189 + H^+$) in the mass spec and the same 1H -NMR resonances as the product obtained from fermentation.

The biosynthetic pathway resulting in HMBPA production is set forth in Figure 2. The chemical structure of [R]-3-(2-hydroxy-3-methyl-butrylamino)-propionic acid (HMBPA) is depicted in Figure 3. As depicted in Figure 2, HMBPA is the condensation product of α -hydroxyisovaleric acid (α -HIV) and β -alanine, catalyzed by the PanC enzyme. α -HIV is generated by reduction of α -KIV, a reaction which is catalyzed by the reductases PanE (*e.g.*, PanE1 and/or PanE2) and/or IlvC.

Based on the chemical structure and biosynthetic pathway leading to HMBPA production, the present inventors formulated the following model to describe the interaction between the previously known pantothenate and isoleucine-valine (*ilv*) pathways and the newly characterized HMBPA biosynthetic pathway. In at least one aspect, the model states that there exist at least two pathways in microorganisms that compete for α -KIV, the substrate for the biosynthetic enzyme PanB, namely the pantothenate biosynthetic pathway and the HMBPA biosynthetic pathway. (A third and fourth pathway competing for α -KIV are those resulting in the production of valine or leucine from α -KIV, see *e.g.*, Figure 1). At least the pantothenate biosynthetic pathway and the HMBPA biosynthetic pathway further produce competitive substrates for the enzyme PanC, namely α -HIV and pantoate. The model predicts that reducing PanB activity will increase α -KIV availability for α -HIV synthesis (and ultimately, HMBPA

synthesis) and decrease the amount of pantoate and/or pantothenate synthesized by a microorganism. Conversely, increasing PanB activity will increase pantoate and ketopantoate availability for pantoate/pantothenate synthesis. The following examples provide experimental support for the model and further exemplify processes for increasing the production of HMBPA based on the model.

EXAMPLES II-VI:

For Examples II-VI, quantitation of pantothenate and/or HMBPA was performed as follows. Aliquots of fermentation media were diluted 1:100 and aliquots of test tube cultures were diluted 1:10 in water or 5% acetonitrile prior to injection on a Phenomenex Aqua™ 5μ C18 HPLC column (250 x 4.60mm, 125A). Mobile phases were A = 5% acetonitrile, 50 mM monosodium phosphate buffer adjusted to pH 2.5 with phosphoric acid; and B = 95% acetonitrile, 5% H₂O.

Linear gradients were as follows.

Minutes	Solvent A	Solvent B
0	100%	0%
16	100%	0%
17	0%	100%
20	0%	100%
21	100%	0%

Additional parameters and apparatus were as follows: Flow rate = 1.0 ml/min; Injection volume = 20 μl; Detector = Hewlett Packard 1090 series DAD UV detector-3014, Signal A = 197 nm, ref. = 450 nm, Firmware revision E; Column heater = Oven temperature 40°C; Hardware = Hewlett Packard Kayak™ XA; and Software = Hewlett Packard Chemstation Plus™ family revision A.06.03[509].

Under these fermentation conditions, PA824 routinely yields approximately 30-40 g/L pantothenate. HMBPA elutes at approximately 13 minutes in this system.

Example II: Ketopantoate Reductase Contributes to the Production of HMBPA and Increasing Ketopantoate Reductase Activity in *Bacillus* Results in Enhanced HMBPA Production

5 As described in Example I, a novel HPLC peak corresponding to HMBPA was observed in microorganisms overexpressing *panE1* indicating that increased ketopantoate reductase contributes to the production of HMBPA (in addition to production of pantothenate). As mentioned previously, two genes in the *B. subtilis* genome are homologous to the *E. coli panE* gene encoding ketopantoate reductase and
10 have been named *panE1* and *panE2*. In *Bacillus*, the *panE1* gene encodes the major ketopantoate reductase involved in pantothenate production, while *panE2* does not contribute to pantothenate synthesis. In fact, overexpression of *panE2* from a *P₂₆* promoter leads to a reduction in pantothenate titer (see *e.g.*, U.S. Patent Application Serial No. 09/400,494).

15 Accordingly, it was tested whether, beside being produced by the *panE1* gene product, it was possible that a significant portion of the α -HIV necessary to make HMBPA was being produced by the *panE2* gene product. It was hypothesized that the *panE2* gene product is an enzyme that can reduce α -KIV to α -HIV, but that can not significantly reduce ketopantoate to pantoate.

20 To test the hypothesis, *panE2* was deleted from pantothenate production strain PA824 (described in Example I) by transforming with a Δ *panE2::cat* cassette from chromosomal DNA of strain PA248 (Δ *panE2::cat*) (set forth as SEQ ID NO:24, for construction see *e.g.*, U.S. Patent Application Serial No. 09/400,494) to give strain PA919. Three isolates of PA919 were compared to PA824 for pantothenate and
25 HMBPA production in test tube cultures grown in SVY plus β -alanine.

Table 1. Production of pantothenate and HMBPA by derivatives of PA824 and PA880 grown at 43°C in 48 hour test tube cultures of SVY glucose + β -alanine⁵.

5					
Strain	new trait	parent	OD ₆₀₀	[pan] g/l	[HMBPA] g/l
PA824	—		13.9	4.3	0.64
PA919-1	Δ panE2::cat	PA824	13.2	4.2	0.15
PA919-2	"	"	14.8	3.8	0.13
PA919-3	"	"	18.0	5.5	0.14

As indicated by the data in Table 1, all three isolates of PA919 produced about four-fold lower HMBPA than PA824 demonstrating that the *panE2* gene product is a potent contributor to HMBPA synthesis. Moreover, significant increases in

- 5 HMBPA production can be achieved simply by overexpression of *panE2*. An exemplary plasmid for the overexpression of *panE2*, named pAN238, is set forth as SEQ ID NO:25 (Figure 10).

10 **Example III. Increasing Production of HMBPA by Reducing PanB Activity in Microorganisms.**

- Strains derived from PA365 (the RL-1 lineage equivalent of PA377, described in U.S. Patent Application Serial No. 09/667,569) which are deleted for the *P₂₆ panBCD* cassette and which contain a *P₂₆ panC*D* cassette amplified at the *vpr* locus and either the wild type *P₂₆ panB* cassette (PA666) or a *P₂₆ ΔpanB* cassette (PA664) amplified at the *bpr* locus were constructed as follows. An alignment of the C-terminal amino acids of known or suspected PanB proteins is shown in Figure 6. Three regions called 1, 2 and 3, that were identified having conserved or semi-conserved amino acid residues, are indicated by arrows at the top of the figure. The *B. subtilis* PanB protein (RBS02239) is underlined. Two of the PanB proteins (RCY14036 and CAB56202.1) are missing region 3 while the latter PanB protein is also missing region 2 and has non-conserved amino acid residues occupying region 1.

- B. subtilis* PanB variants were created that were missing regions 1, 2 and 3. The desired variants were created by designing 3' PCR primers to amplify the *B. subtilis pan B* gene such that region 3, regions 2 and 3, or all three regions would be missing from the final product. The PCR products were generated and cloned into *E. coli* expression vector pASK-1BA3, creating plasmids pAN446, pAN447, and pAN448, respectively. The plasmids were then transformed into *E. coli* strain SJ2 that contains the *panB6* mutation to test for complementation. Only pAN446, which is missing region 3, was able to complement. This indicates that region 3 is not essential for *B. subtilis* PanB activity but that region 2 is required for activity or stability.

The next step in this analysis was to transfer the *panB* gene from pAN446 to a *B. subtilis* expression vector and then introduce it into a strain appropriate for testing activity of the encoded PanB protein in *B. subtilis*. To do this, a strain that is deleted for the *P*₂₆ *panBCD* operon was first created. This was accomplished by first inserting a *cat* gene between the *Bse*RI site located just upstream of the *panB* RBS and the *Bgl*/II site located in *panD*, creating plasmid pAN624 (*Figure 7*). The sequence of pAN624 is set forth as SEQ ID NO:20. The resulting deletion-substitution mutation (Δ *panBCD::cat624*), which removes all of *panB* and *panC*, was crossed into PA354 by transformation, with selection for resistance to chloramphenicol on plates supplemented with 1 mM pantothenate. One of the transformants was saved and named PA644. Chromosomal DNA isolated from PA644 was analyzed by PCR and was shown to contain the deletion-substitution mutation. As expected, PA644 requires pantothenate for growth but retains the engineered *ilv* genes (*P*₂₆*ilvBNC* *P*₂₆*ilvD*) as well as the *P*₂₆*panE1* gene originally present in PA354. Thus, it has all the enzymes involved in pantoate synthesis overproduced except PanB. The gene containing the shortest *panB* deletion was inserted into *B. subtilis* expression vector pOTP61 (described in US patent application Serial No. 09/667,569), creating plasmid pAN627. At the same time, a wild-type *panB* control gene was inserted into pOTP61, creating plasmid pAN630. The *Not*I fragments of each plasmid, lacking *E. coli* vector sequences, were ligated and transformed into PA644, with selection for resistance to tetracycline.

One transformant from each transformation was saved and further transformed with chromosomal DNA from PA628 with selection for Pan⁺. PA628 contains a multicopy *P*₂₆*panC*D* expression plasmid (pAN620) integrated at the *vpr* locus. In order to determine the effects of the *panB* gene mutation directly on pantothenate production, plasmid pAN620, set forth as SEQ ID NO:21 and illustrated schematically in *Figure 8*, provides the remaining two enzymes required for pantothenate synthesis (PanC and PanD). Four transformants from each transformation were isolated, grown in SVY medium containing 10 g/L aspartate for 48 hours, then assayed for pantothenate production. Transformants with the 3' deleted *panB* gene were named PA664 and those containing the wild-type gene were called PA666. The data showed that the 3' deleted *panB* gene in PA664 encodes a PanB protein with greatly reduced activity. To test for HMBPA production, test tube cultures of PA365, PA666, and PA664 were grown in SVY + aspartate medium with and without added α -KIV or

pantoate for 48 hours and then assayed for HMBPA and pantothenate as described previously.

Table 2. Effect of PanB activity and addition of precursors on HMBPA and pantothenate production, 48 hour test tube culture data, SVY + aspartate (10 g/L) medium.

Strain	pan operon		panC*D plasmid		panB plasmid		no additions		+ α -KIV (5 g/L)		+ pantoate (5 g/L)	
PA365	P_{26}	panBCD	NONE	NONE	NONE		[pan] (g/L)	HMBPA peak*	[pan] (g/L)	HMBPA peak	[pan] (g/L)	HMBPA peak
PA666	Δ panBCD::cat		pAN620	pAN630			3.0	0.71	3.2	1.28	4.8	0.38
PA664	Δ panBCD::cat		pAN620	pAN627			3.7	0.55	3.3	1.70	5.2	0.26
							0.3	1.39	0.6	1.76	2.5	0.74

* HMBPA peak = peak area $\times 10^{-3}$

The data presented in Table 2 demonstrate that in the absence of supplements, PA664 produced the most HMBPA while PA666 produced the least, indicating an inverse correlation between PanB activity and HMBPA production. This is consistent with the model which predicts that the two pathways compete for α -KIV, the substrate for PanB, and produce competitive substrates for PanC; lowering PanB activity would be expected to increase α -KIV availability for α -HIV synthesis and increase HMBPA production, correspondingly decreasing the amount of pantoate synthesized. When α -KIV is added to the medium, all three strains produced significantly more HMBPA. This result evidences that α -KIV is a precursor to HMBPA, as described in Figure 2, and that excess α -KIV favors HMBPA production. This result also suggests that synthesis of HMBPA is at least partially due to an overflow effect of excess α -KIV production. When pantoate was added to the medium, HMBPA was reduced by roughly 50 percent in all three strains. Conversely, the strains each produced significantly more pantothenate. This result is also consistent with the model that the two pathways produce competing substrates for PanC (α -HIV and pantoate). Taken together, the above results further indicate that decreasing pantoate synthesis should be beneficial in promoting HMBPA production as well as reducing pantothenate levels.

Example IV. Methods for Regulating HMBPA:Pantothenate Levels

As demonstrated in Examples I and II, PanE1 and/or PanE2 contribute to enhanced HMBPA production as does reduced PanB activity. This Example demonstrates that overexpressing PanE1 increases HMBPA production relative to pantothenate production whereas overexpressing PanB decreases HMBPA production relative to pantothenate production. Furthermore, in strains overexpressing IlvC, HMBPA production is enhanced.

PA668 is a derivative of PA824 that contains extra copies of *P*₂₆ *panB* amplified at the *vpr* or *panB* locus. PA668 was constructed using the *panB* expression vector (pAN636) which allows for selection of multiple copies using chloramphenicol. The sequence of pAN636 is set forth as SEQ ID NO:22 and the vector is depicted schematically in Figure 9. The pAN636 *NotI* restriction fragment, missing the *E. coli* vector sequences, was ligated and then used to transform PA824 with selection on plates containing 5 μ g/ml chloramphenicol. Transformants resistant to 30 μ g/ml

chloramphenicol were isolated and screened for pantothenate production in 48 hour test tube cultures. The isolates shown produce less HMBPA than PA824 (conversely producing about 10 percent more pantothenate than PA824). A second strain, called PA669, was constructed which is PA824 with extra copies of *P26 panE1* amplified at the *vpr* or *panE1* locus. Strain PA669 was constructed by transforming PA824 with the self-ligated *NotI* fragment of plasmid pAN637 with selection for resistance to chloramphenicol. The sequence of pAN637 is set forth as SEQ ID NO:23 and the vector is depicted schematically in Figure 10. Two isolates of PA669 were chosen for further study; PA669-5 produces less PanE1 than PA669-7 as judged by SDS-PAGE analysis of total cell extracts made from the two strains.

Test tube cultures of strains PA824, PA668-2, PA668-24, and the two isolates of PA669 (PA669-5 and PA669-7) were grown in three different media (SVY, SVY + aspartate, and SVY + aspartate + pantoate) for 48 hours and then assayed for pantothenate, HMBPA, and β -alanine (Table 3).

Table 3. Effect of extra copies of *panB* and *panE1* on pantothenate and HMBPA production by PA824, 48 hour test tube culture data, SVY medium.

Strain	<i>panB</i> plasmid	<i>panE</i> plasmid	no additions			+ aspartate (10 g/L)			+ aspartate (10 g/L) & pantoate (5 g/L)		
			[pan] (g/L)	[β-ala] (g/L)	HMBPA *	[pan] (g/L)	[β-ala] (g/L)	HMBPA	[pan] (g/L)	[β-ala] (g/L)	HMBPA
PA824	NONE	NONE	1.8	0.05	<0.1	4.7	2.5	0.53	5.6	2.5	<0.10
PA668-2	pAN636	NONE	1.5	<0.04	<0.1	5.0	1.6	<0.10	4.9	1.2	<0.10
PA668-24	pAN636	NONE	1.8	0.05	<0.1	4.9	2.8	0.34	6.1	2.6	<0.10
PA669-5	NONE	pAN637	1.8	0.04	<0.1	4.2	3.1	0.74	5.8	2.6	0.30
PA669-7	NONE	pAN637	1.8	0.06	<0.1	3.7	3.2	1.41	5.2	2.5	0.75

5 * HMBPA = peak area X 10⁻³

None of the strains produced detectable quantities of HMBPA in SVY medium. All strains produced roughly equivalent amounts of pantothenate and low amounts of β -alanine indicating that β -alanine is limiting for both pantothenate and HMBPA synthesis in these cultures and that β -alanine is a precursor for both compounds. When grown in SVY + aspartate medium, the two PA669 isolates produced more HMBPA than PA824 whereas both PA668 isolates produced less HMBPA than PA824. It is noteworthy that the strain that produces the most PanE1 (PA669-7) produced the most HMBPA (and the least pantothenate). This suggests that high levels of PanE1 favor the production of HMBPA at the expense of lower pantothenate synthesis. It is also interesting that PA668-24 produced more HMBPA than PA668-2, even though SDS-PAGE analysis of extracts from the two strains showed that they produce roughly equivalent levels of PanB. The SDS-PAGE analysis also showed that PA668-24 makes much more IlvC than PA668-2. Based on these data, it is proposed that IlvC influences HMBPA synthesis by increasing steady state levels of α -KIV and/or by catalyzing α -HIV formation from α -KIV, thereby accounting for the observed shift towards production of HMBPA.

The final set of data in Table 3 shows that adding pantoate to the growth medium decreased HMBPA production by all strains that had previously produced detectable levels, *e.g.*, by shifting synthesis towards pantothenate. This further supports the model that α -HIV and pantoate are competitive substrates for PanC.

Example V: Increasing HMBPA Production by Limiting Serine Availability

It was hypothesized that the ratio of pantothenate to HMBPA production could also be controlled by regulating the availability of serine or methylene tetrahydrofolate in the microorganism cultures. In particular, it is proposed that decreasing the availability of serine could increase HMBPA production relative to pantothenate production, whereas increasing the availability of serine would decrease the production of HMBPA relative to pantothenate production. This method is based on the understanding that the PanB substrate, methylenetetrahydrofolate is derived from serine. Thus, regulating serine levels should effectively regulate PanB substrate levels. To test this hypothesis, PA824 was grown in test tube cultures of SVY glucose plus 5 g/L β -alanine and \pm 5 g/L serine for 48 hours at 43°C.

Table 4: *Production of HMBPA and pantothenate by PA824 with and without the addition of serine*

serine added at 5 g/L	OD ₆₀₀	[pan] g/L	[HMBPA] g/L
-	16.3	4.9	0.84
-	14.0	4.5	0.80
+	13.1	6.4	0.56
+	12.9	6.0	0.62

5

As demonstrated in Table 4, addition of serine decreases the level of production while conversely increasing pantothenate production. At least one method of decreasing methylene tetrahydrofolate levels in order to regulate HMBPA production levels is to decrease the activity of serine hydroxymethyl transferase (the *glyA* gene product), thereby decreasing methylene tetrahydrofolate biosynthesis in appropriately engineered microorganisms. At least one method of decreasing serine levels in order to regulate HMBPA production is to decrease the activity of 3-phosphoglycerate dehydrogenase (the *serA* gene product).

15

Example VI: Increasing HMBPA Production by Modifying Culture Conditions for Recombinant Microorganisms

In at least one fermentation (Fermentation P162), levels of HMBPA production reached 35 g/L. Briefly, fermentation of strain PA824 was carried out as described in Example I but utilizing PFM-155 medium formulated as follows.

20

BATCH		
	MATERIAL	g/L (final)
1	Amberex 1003	5
2	Cargill 200/20 (soy flour)	40
3	Na Glutamate	5
4	(NH ₄) ₂ SO ₄	8
5	MgSO ₄ ·7H ₂ O	1

6	MAZU DF204C	1
7	H ₂ O	qs to 4 L

Added After Sterilization and Cool Down

1	KH ₂ PO ₄	10
2	K ₂ HPO ₄ ·3H ₂ O	20
3	H ₂ O	qs to 400 ml
1	80% Glucose	20
2	CaCl ₂ ·2H ₂ O	0.1
1	Sodium Citrate	1
2	FeSO ₄ ·7H ₂ O	0.01
3	SM-1000X	1 X

FEED

	MATERIAL	g/L (final)
1	80% Glucose	800
2	CaCl ₂ ·2H ₂ O	0.8
3	H ₂ O	qs to 3500 ml

Added After Sterilization and Cool Down

1	Sodium Citrate	2.0
2	FeSO ₄ ·7H ₂ O	0.02
3	SM-1000X	2 X
4	Glutamate Na	5.0
5	H ₂ O	qs to 500 ml

5

However, as a result of loss of process control during the fermentation, the dissolved oxygen became limiting between 16 and 17 hours and glucose began to accumulate after 16 hours.

These changes in fermentation conditions produced the following significant results at or after 16 hours. Namely, synthesis of HMBPA began to increase with a corresponding decrease in pantothenate synthesis. In the four hour interval before 16 hours the culture produced 7 g/l HMBPA, four hours afterwards, 9.0 g/l.

- 5 Pantothenate was the reverse with 10 g/l and 6.0 g/l produced between 12-16 hours and 16-20 hours, respectively. Between 20 and 36 hours the average rate of HMBPA synthesis was 1.0 g/l hr. Overall, fermentation P162 produced 35 g/l of HMBPA in 36 hours.

- 10 Thus, it appears that overfeeding of glucose, and/or limitation of dissolved oxygen (e.g., beginning at about 16 hours) leads to an increase in HMBPA production. Accordingly, two methods for increasing HMBPA production (relative to pantothenate production) are to increase steady state glucose levels and/or decrease steady state dissolved oxygen levels.

15

- Equivalents Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the
- 20 following claims.

What is claimed:

1. A process for the production of 3-(2-hydroxy-3-methyl-butrylamino)-propionic acid (HMBPA), comprising culturing a microorganism under conditions such that HMBPA is produced and detecting the HMBPA produced by said microorganism.
2. A process for the production of 3-(2-hydroxy-3-methyl-butrylamino)-propionic acid (HMBPA), comprising culturing a microorganism under conditions such that HMBPA is produced and isolating the HMBPA produced by said microorganism.
3. A process for the production of 3-(2-hydroxy-3-methyl-butrylamino)-propionic acid (HMBPA), comprising culturing a microorganism having increased keto reductase activity or increased pantothenate synthetase activity in the presence of excess α -ketoisovalerate and excess β -alanine, such that HMBPA is produced.
4. A process for the production of 3-(2-hydroxy-3-methyl-butrylamino)-propionic acid (HMBPA), comprising culturing a microorganism having increased keto reductase activity and increased pantothenate synthetase activity in the presence of excess α -ketoisovalerate and excess β -alanine, such that HMBPA is produced.
5. The process of claim 3 or 4, wherein said microorganism comprises a modified *panE* gene.
6. The process of claim 5, wherein the *panE* gene is overexpressed, deregulated or present in multiple copies.
7. The process of claim 3 or 4, wherein said microorganism comprises a modified *panE1* gene.
8. The process of claim 3 or 4, wherein said microorganism comprises a modified *panE2* gene.

9. The process of claim 3 or 4, wherein said microorganism comprises a modified *panE1* gene and a modified *panE2* gene.

5 10. The process of claim 3 or 4, wherein said microorganism comprises a modified *panC* gene.

11. The process of claim 3 or 4, wherein the *panC* gene is overexpressed, deregulated or present in multiple copies.

10

12. The process of claim 3 or 4, wherein said microorganism further has increased acetohydroxyacid isomeroreductase activity.

13. A process for the production of 3-(2-hydroxy-3-methyl-
15 butyrylamino)-propionic acid (HMBPA), comprising culturing a microorganism having increased acetohydroxyacid isomeroreductase activity in the presence of excess α -ketoisovalerate and excess β -alanine, such that HMBPA is produced.

14. The process of claim 12 or 13, wherein said microorganism
20 comprises a modified *ilvC* gene.

15. The process of claim 14, wherein the *ilvC* gene is overexpressed, deregulated or present in multiple copies.

25 16. The process of any one of claims 3, 4 or 11, wherein said microorganism further has reduced ketopantoate hydroxymethyltransferase activity.

17. The process of claim 16, wherein said microorganism comprises a modified *panB* gene.

30

18. The process of claim 16, wherein said microorganism has been deleted for the *panB* gene.

19. A process for the production of 3-(2-hydroxy-3-methyl-butrylamino)-propionic acid (HMBPA), comprising culturing a microorganism having reduced ketopantoate hydroxymethyltransferase activity in the presence of excess α -ketoisovalerate and excess β -alanine, such that HMBPA is produced.
20. A method for enhancing production of 3-(2-hydroxy-3-methyl-butrylamino)-propionic acid (HMBPA) relative to pantothenate, comprising culturing a recombinant microorganism under conditions such that the HMBPA production is enhanced relative to pantothenate production.
21. A process for the production of 2-hydroxyisovaleric acid (α -HIV), comprising culturing a microorganism which overexpresses PanE1 or PanE2 and which further has reduced PanC or PanD activity under conditions such that α -HIV is produced.
22. A process for the production of 3-(2-hydroxy-3-methyl-butrylamino)-propionic acid (HMBPA), comprising culturing a recombinant microorganism having decreased expression or activity of *serA* or *glyA* under conditions such that HMBPA is produced.
23. A process for the production of 3-(2-hydroxy-3-methyl-butrylamino)-propionic acid (HMBPA), comprising culturing a recombinant microorganism having decreased expression or activity of *serA* and *glyA* under conditions such that HMBPA is produced.
24. The process of any one of the preceding claims wherein the microorganism is cultured under conditions of increased steady state glucose.
25. The process of any one of the preceding claims wherein the microorganism is cultured under conditions of decreased steady state dissolved oxygen.
26. The process of any one of the preceding claims wherein the microorganism is cultured under conditions of decreased serine.
27. A product produced according to any one of the above claims.

28. A recombinant microorganism that produces 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA), the microorganism having a modification in at least one gene encoding ketopantoate reductase that results in increased reductase activity and having a mutation or deletion in the *panB* gene that results in reduced ketopantoate hydroxymethyltransferase activity.

29. The recombinant microorganism of claim 28, wherein the gene encoding ketopantoate reductase is a *panE* gene.

30. The recombinant microorganism of claim 29, wherein the *panE* gene is *panE1*.

31. The recombinant microorganism of claim 29, wherein the *panE* gene is *panE2*.

32. The recombinant microorganism of claim 28, wherein the microorganism has a modification in *panE1* and *panE2*.

33. The recombinant microorganism of claim 28, further having a modification in *ilvC* that results in increased acetohydroxyacid isomeroreductase activity.

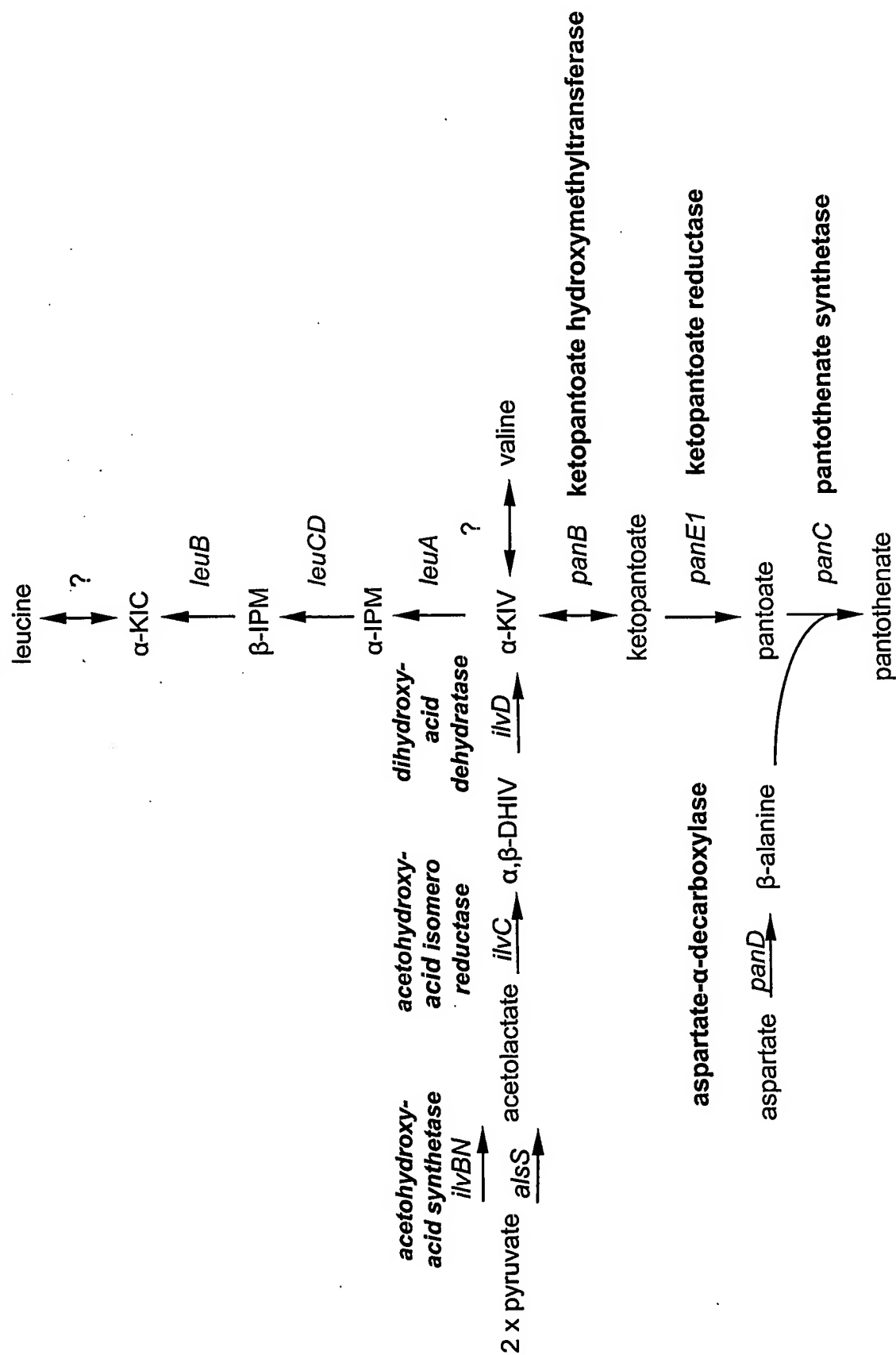
34. A recombinant microorganism that produces 3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid (HMBPA), the microorganism having a modification in *ilvC* that results in increased acetohydroxyacid isomeroreductase activity and having a mutation or deletion in the *panB* gene that results in reduced ketopantoate hydroxymethyltransferase activity.

35. The recombinant microorganism of any one of claims 28 to 34, wherein said microorganism belongs to the genus *Bacillus*.

36. The recombinant microorganism of claim 35, which is *Bacillus subtilis*.

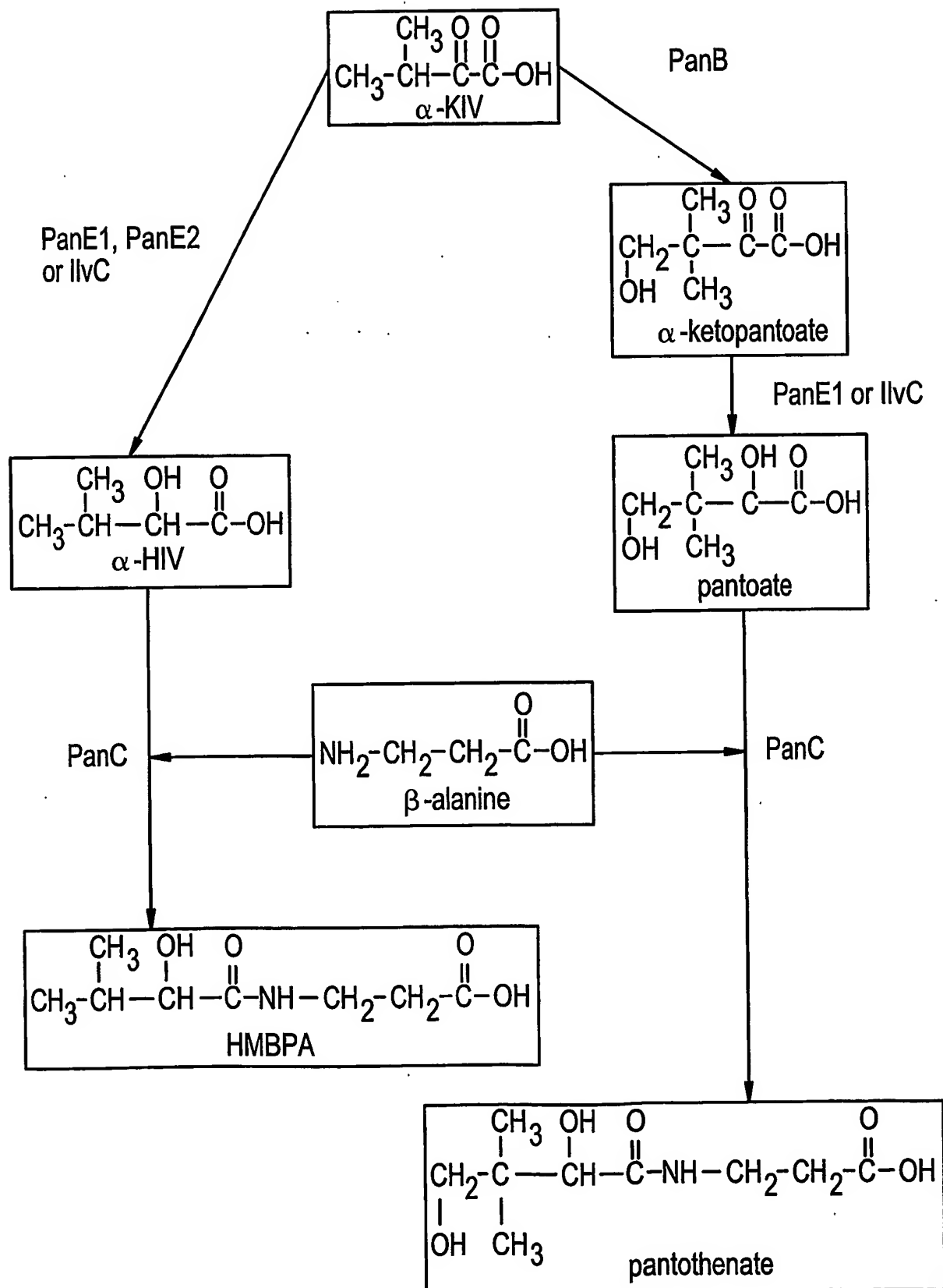
1/11

FIG. 1



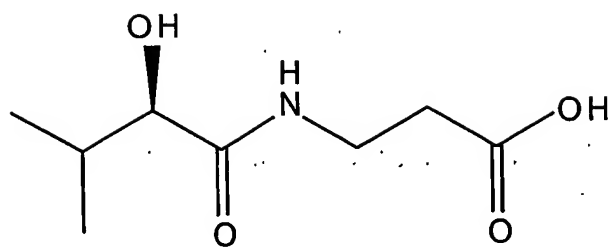
2/11

FIG. 2



3/11

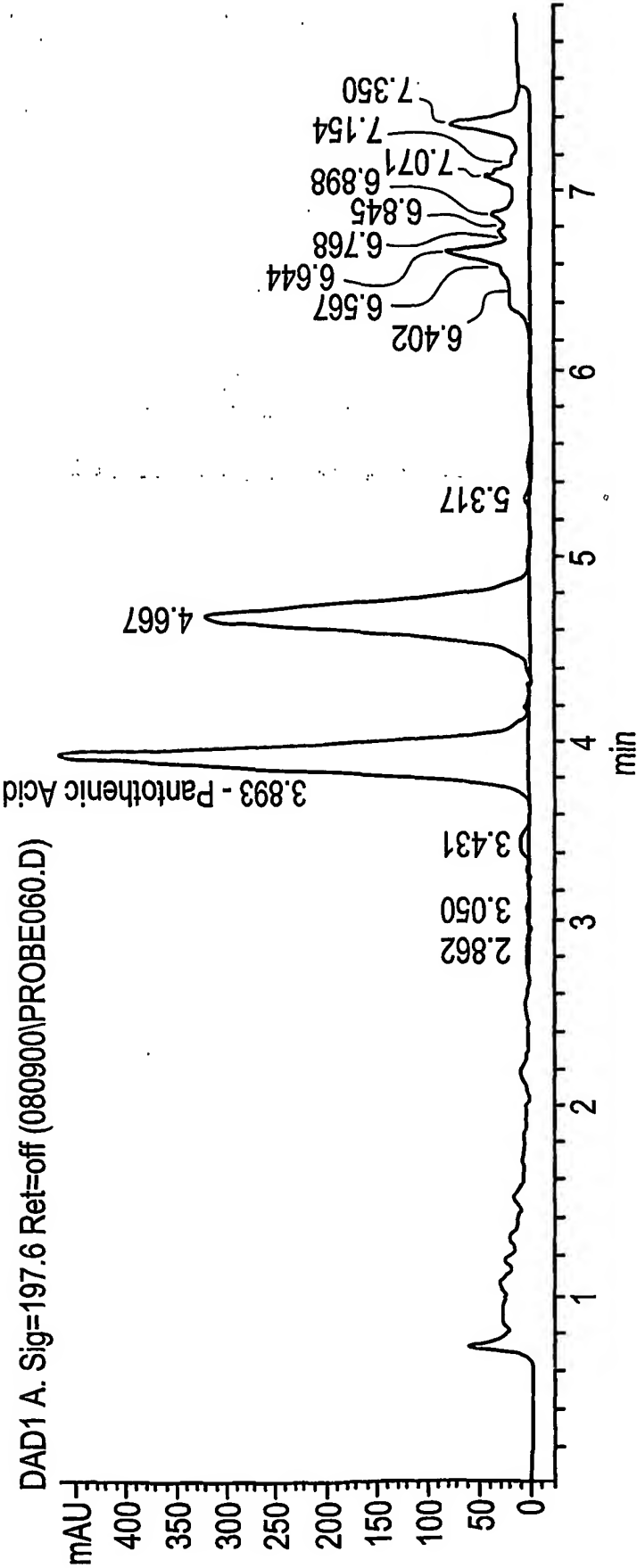
FIG. 3



[R]-3-(2-hydroxy-3-methyl-butyrylamino)-propionic acid ("HMBPA")

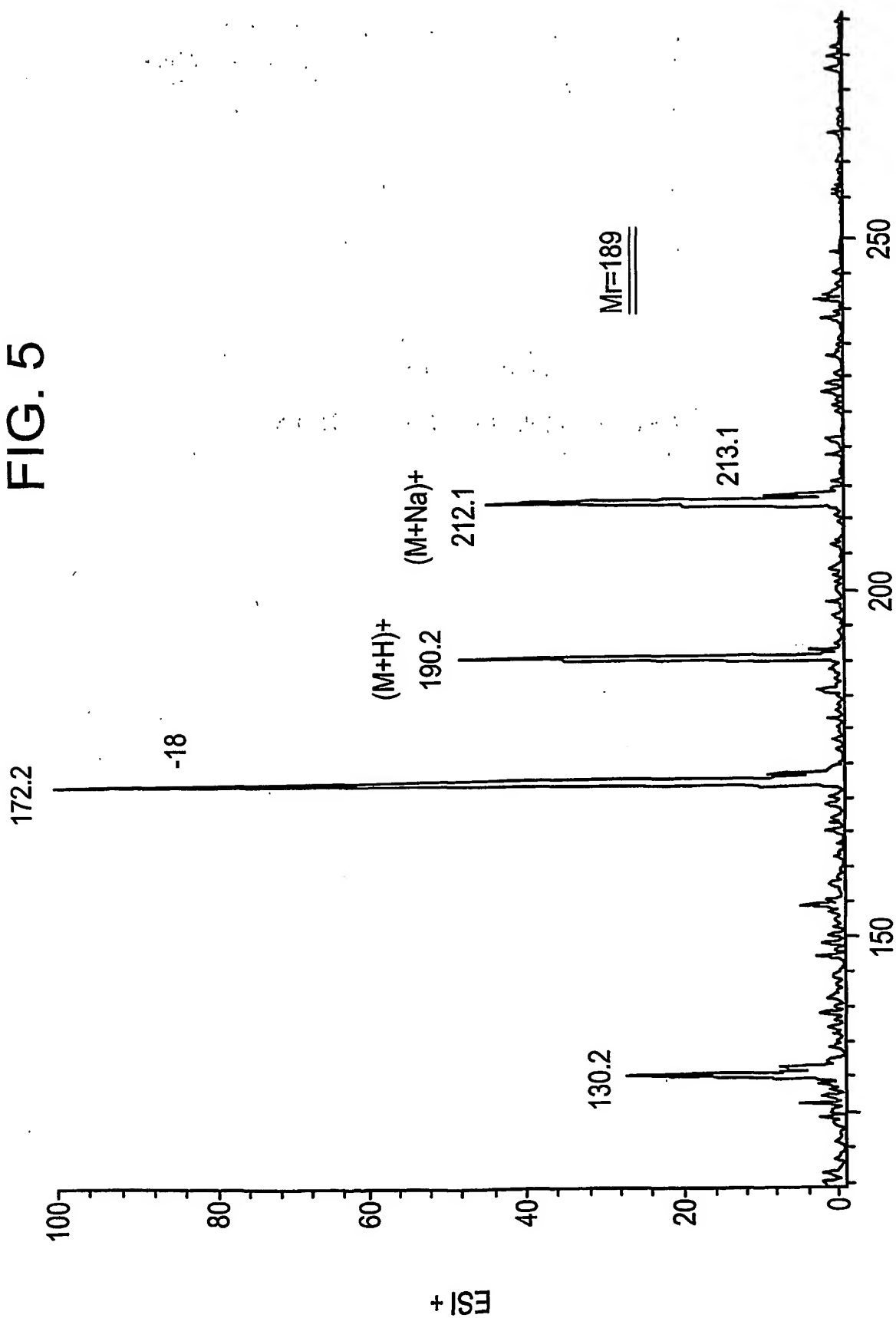
11/4

FIG. 4



5/11

FIG. 5



6/11

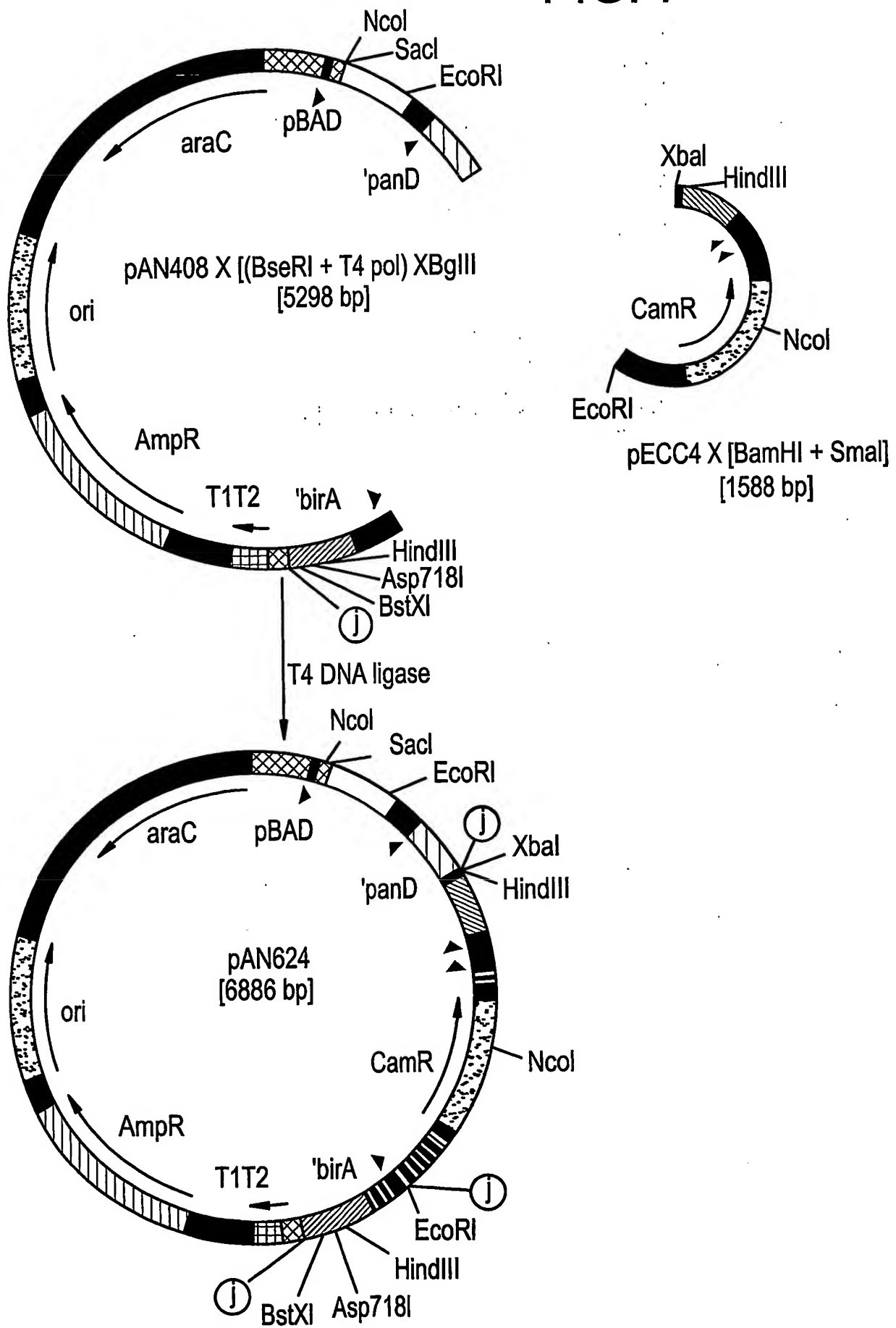
FIG. 6

An alignment of the C-terminal amino acids from
all known or suspected PanB proteins.

	1	2	3	
	↓	↓	↓	
CAA10222.1	AYVAEV	KGVTF	FGA	EHGFSA-----
APE0676	RYAEDV	RNGRF	FPGE	EHVVH--AKEPLEDIS-----
RYP04152	LYIEEV	DGGIY	PAE	EHTFQ-----
REC04341	QYMAEV	ESGVY	PGE	EHSEH-----
sp Q09672	RYIYEV	EQGLY	PAE	EHSTQ-----
RNG01188	AYVAEV	KAKTF	FPAA	EHIFAD-----
RNM00107	AYVAEV	KAKTF	FPAA	EHIFAD-----
RPA02174	AYVRAV	KDVSF	FPAA	EHGFNA-----
RCA00999	SYAKEV	REGTF	FPDE	AHSFK-IDQSIIDEITK-----
REF01843	KYIEEV	KDGVF	FPGP	EHGFK-ISDDVLEKLY-----
RCJ02253	KYRDEV	KSGIF	FP SQ	EHSTFDYLDDELLDKLY-----
RBS02239	GYVQDV	RHRAF	PEQ	KHSFQ-MNQTVLDGLYGGK-----
RDR03436	HYAAEV	RAREF	FP SK	DNSEFV-MKDEVLDKLY-----
RCY14036	KFSGEV	RQRQF	FP RR	G-----
RPA08114	RFAEDV	RERRF	PEA	RHCFAMRE-----
RRC02991	AYAAEV	RSRAFP	PAP	EFSEFDEVKK-----
SCC75A.02	AYAEDV	VGGTF	FPAD	EHSVH-----
CAA65397.1	AYIADI	HAGTF	FPGE	AESF-----
CAB56202.1	EWVAAE	KLN	---	-----
RMT01063	QYAEV	VAGGV	FPAD	EHSE-----
sp Q10505	QYAEV	VAGGV	FPAD	EHSE-----
RML00370	QYAEV	VASAV	FP AE	EHCF-----
RML00622	QYAEV	VASAV	FP AE	EHCF-----
RAA01082	NFKIDV	EGGNF	FP SE	EESYG-----
RHP00462	QYADV	KKGNF	FP NE	LESYH-----
RPG00121	HYIADV	KSMDF	FP NK	DEQY-----
RPH00184	TFREEV	KEGKF	FP GR	EHYWEFQDKEEFKRIKDNVMKKLNL-----
RPF00767	EFRKEV	KEGKF	FP GK	EHYWEYQDKETFNRIKENVMRKLRL-----
TM1728	EFRREV	KKGLF	FP TE	EHSTDKSKGGVSS-----
PAB0570	EFKKDV	KEGKF	FP GR	EHYWEFQDKEEFKRIKESVLRKVD-----
RSC03568	EYIASV	EDRTF	FP ER	GTHTFKVKEDLWNEFLSSINEK-----
AAD37248.1	QYREV	KSRAY	PAE	QHTYPIPKHEELVEFQKAVDELPEEK-----
T3F17.24	EYKEEV	SKKV	FP GP	SHSPYKITASELDGFLTELQKLGFDKAASAAALAAENMEPSK-----
	:	*	:	*

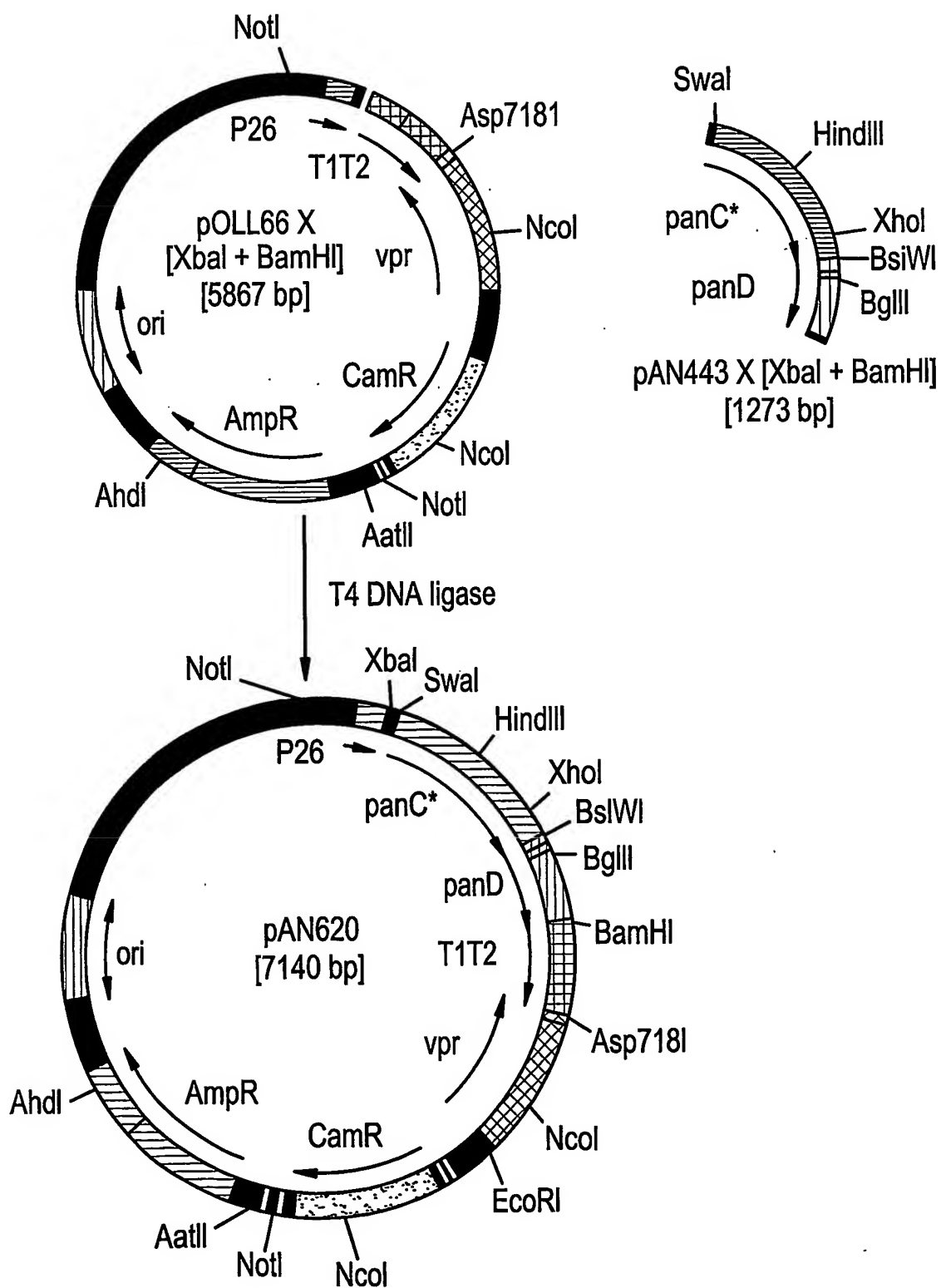
7/11

FIG. 7



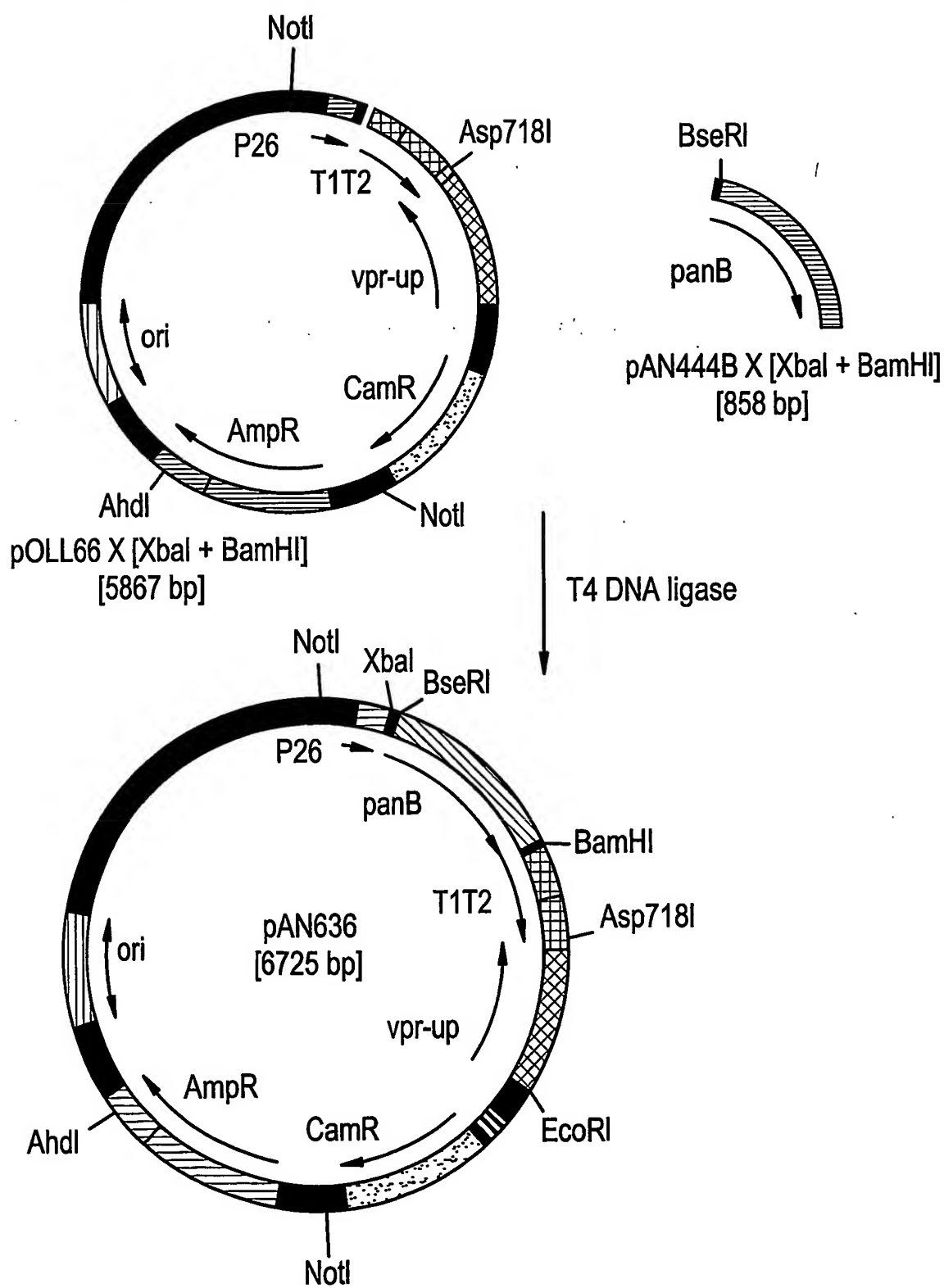
8/11

FIG. 8



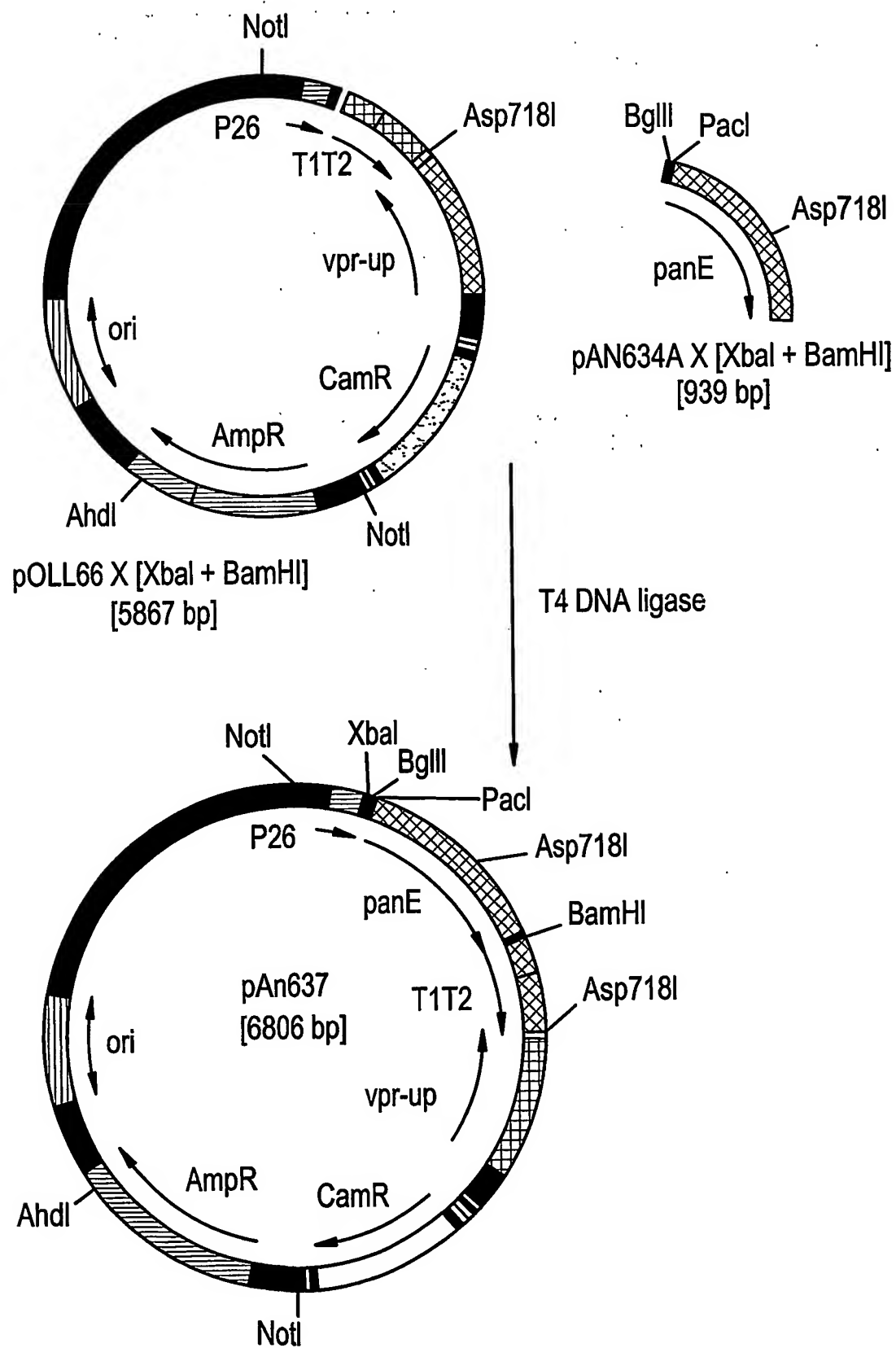
9/11

FIG. 9

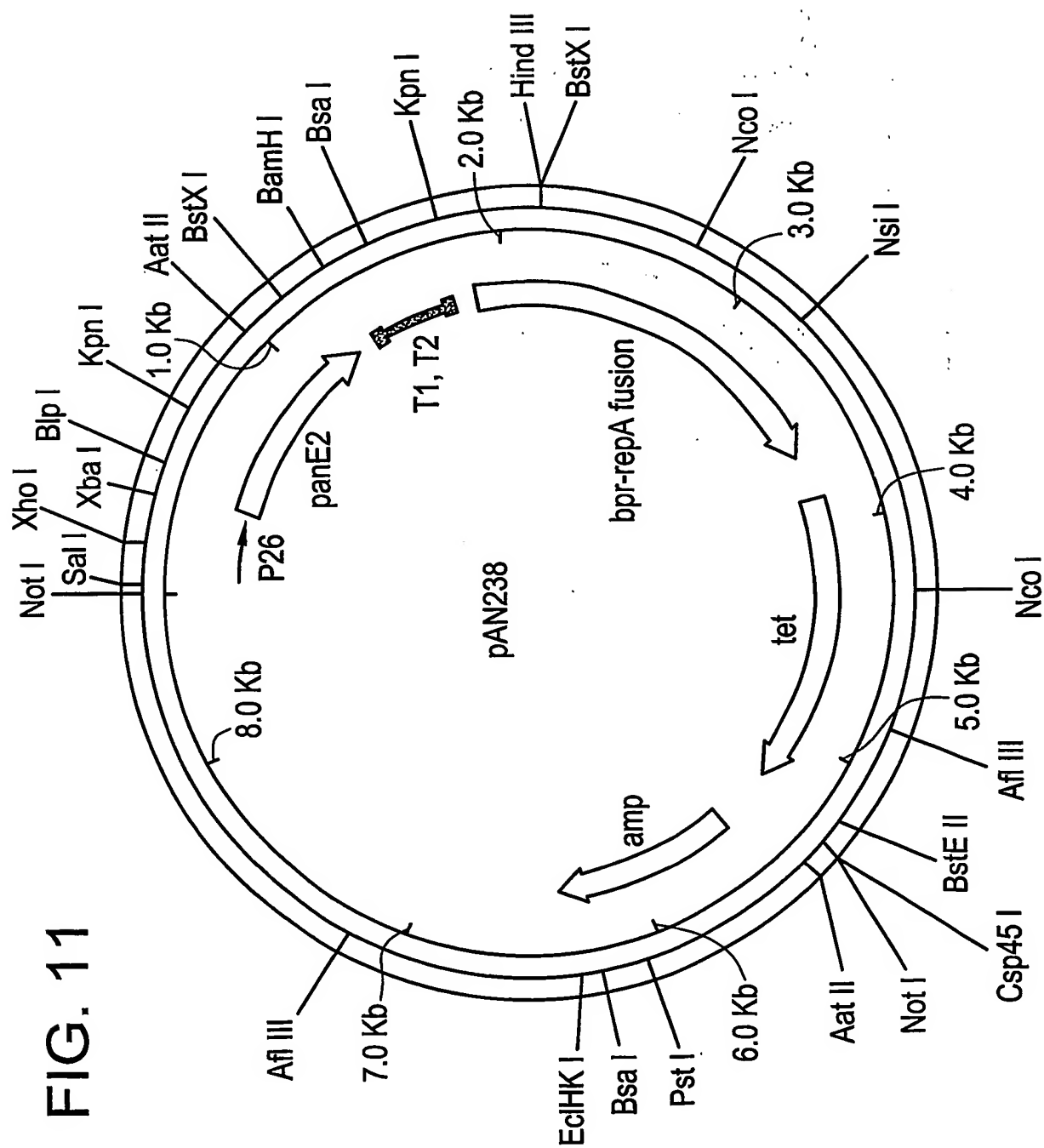


10/11

FIG. 10



11/11



SEQUENCE LISTING

<110> OmniGene BioProducts Inc., *et al.*

<120> METHODS AND MICROORGANISMS FOR THE PRODUCTION OF 3-(2-HYDROXY-3-METHYL-BUTYRLAMINO)-PROPIONIC ACID (HMBPA)

<130> BGI-146PC

<140>

<141>

<160> 25

<170> PatentIn Ver. 2.0

<210> 1

<211> 194

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:promoter sequence

<220>

<221> -35_signal

<222> (136)..(141)

<220>

<221> -10_signal

<222> (159)..(164)

<400> 1

```
gctattgacg acagctatgg ttcaactgtcc accaaccacaaa actgtgtctca gtaccgccaa 60
tattttctccc ttgaggggta caaagagggtg tccctagaag agatccacgc tgtgtataaaa 120
ttttacaaaa aggtattgac tttccctaca ggggtgtgtaa taatttaatt acaggcgggg 180
gcaaccccg cgtgt                                     194
```

<210> 2

<211> 163

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:promoter sequence

<220>

<221> -35_signal

<222> (113)..(118)

<220>

<221> -10_signal

<222> (136)..(141)

<400> 2

gcctacctag cttccaagaa agatataccta acagcacaag agcggaaaga tgttttgttc 60

tacatccaga acaacctctg ctaaaattcc tgaaaaattt tgcaaaaagt tgttgacttt 120

atctacaagg tgtggtataa taatcttaac aacagcagga cgc 163

<210> 3

<211> 127

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:promoter
sequence

<220>

<221> -35_signal

<222> (34)..(39)

<220>

<221> -10_signal

<222> (58)..(63)

<220>

<221> -35_signal

<222> (75)..(80)

<220>

<221> -10_signal

<222> (98)..(103)

<400> 3

gaggaatcat agaattttgt caaaataatt ttattgacaa cgtcttatta acgttgatat 60

aatttaaatt ttatttgaca aaaatgggct cgtgttgtag aataaatgta gtgaggtgga 120

tgcaatg 127

<210> 4

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome
binding site

<400> 4

taaacatgag gaggagaaaa catg 24

<210> 5

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome
binding site

<400> 5

attcgagaaa tggagagaat ataatatg

28

<210> 6

<211> 13

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome
binding site

<400> 6

agaaaggagg tga

13

<210> 7

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome
binding site

<220>

<223> All occurrences of n = any nucleotide

<400> 7

ttaagaaagg aggtgannnn atg

23

<210> 8

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome
binding site

<220>

<223> All occurrences of n = any nucleotide

<400> 8

ttagaaagga ggtgannnnn atg

23

<210> 9

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome
binding site

<220>
<223> All occurrences of n = any nucleotide

<400> 9
agaaaggagg tgannnnnnn atg 23

<210> 10
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:ribosome
binding site

<220>
<223> All occurrences of n = any nucleotide

<400> 10
agaaaggagg tgannnnnna tg 22

<210> 11
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:ribosome
binding site

<400> 11
ccctctagaa ggaggagaaa acatg 25

<210> 12
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:ribosome
binding site

<400> 12
ccctctagag gaggagaaaa catg 24

<210> 13
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:ribosome
binding site

<400> 13
ttagaaagga ggatttaa atg 23

<210> 14
<211> 23
<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome
binding site

<400> 14

ttagaaagga ggtttaatta atg

23

<210> 15

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome
binding site

<400> 15

ttagaaagga ggtgatttaa atg

23

<210> 16

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome
binding site

<400> 16

ttagaaagga ggtgtttaaa atg

23

<210> 17

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome
binding site

<400> 17

attcgagaaa ggaggtgaat ataatatg

28

<210> 18

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:ribosome
binding site

<400> 18

attcgagaaa ggaggtgaat aataatg

27

<210> 19

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: ribosome binding site

<400> 19

attcgtagaa aggaggtgaa ttaatatg

28

<210> 20

<211> 6886

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: vector

<220>

<223> pAN624

<400> 20

```

aagaaaccaa ttgtccatat tgcacagac attgccgtca ctgcgtcttt tactggctct 60
tctcgctaac caaacccgta acccgctta ttaaaagcat tctgtaacaa agcgggacca 120
aagccatgac aaaaacgcgt aacaaaagt tctataatca cggcagaaaa gtccacattg 180
attatttgca cggcgtcaca ctttgctatg ccatagcatt tttatccata agattagcgg 240
atcctacctg acgcttttta tcgcaactct ctactgtttc tccatacccg tttttttggg 300
ctaacaggag gaattaacca tggatccgag ctcgacagta tcaagcactt cacaatctgg 360
gagctgaaag cccgccttat gtagctcata cttgacaaat ccaagggtcaa aatggatatt 420
gtgggcgaca aaataagcgc cgtcaagcaa ttggaatact tcttcagcaa ctgcttcaaa 480
tggctgttca ttctcgacca ttgattaga gattccagta agctgctcaa taaaagcagg 540
gattgattta tttggattaa tgtattttga aaaccgctca gtaatttgtc cattttcgat 600
tacaaccgct gcgatttgta tgattttatc gcctttcttc ggcgaattcc ctgttgtctc 660
tacatctata acaacgaacc gttgcttatt cattaaaaatg gacacctcaa ttcttgcata 720
cgacaaaagt gtaacacggt ttgtacggaa atggagcggc aaaaccgttt tactctcaaa 780
atcttaaaag aaaacccccg ataaaggggg cttttcttct acaaaattgt acgggctggg 840
tcgttcccca gcatttggtc aattttgttt tgatcattca gaacagccac tttcggtcca 900
tggcttgccg cttcttgatc agacatcatt ttgtaggaaa taataatgac cttatctcct 960
tctgcacaa ggcgtgcggc tgcaccgttt aagcatatga cggcgttcc ccgtttacca 1020
ggaataatat acgtttcaag acgtgctcca ttattattat tcacaatttg tactttttca 1080
ttaggaagca ttccacagc atcaatgaga tcctctagag tcgacctgca ggcatgcaag 1140
cttcgcgcga cgtctctcct tatgcgactc ctgcattagg aagcagccca gtagtagggt 1200
gaggccgttg agcaccgccc ccgcaaggaa tggtgcatgc aaggagatgg cgcccaacag 1260
tcccccgccc acggggcctg ccaccatacc caccgcgaaa caagcgtca tgagcccga 1320
gtggcgagcc cgatcttccc catcggtgat gtgcggcagata taggcgccag caaccgcacc 1380
tgtggcgccg gtgatgccgg ccacgatgcg tccggcgtag aggatcaatc ttcattccatt 1440
ccaaggtaaa tcccccttcg ccgtttctgt taccattata caccttttga accttaacgt 1500
aaacgttaag ttttaaaaaa caataaaaaa gacgagcagc atacagcacc cgtctttcac 1560
tttctgtttt aagctaaact tcccgcactt gacagagact ctttttgaag gctttcagaa 1620
agcactcgat acgcgatctg gagctgtaat ataaaaacct tcttcaacta acggggcagg 1680
ttagtacat tagaaaaacc actgtaaaaa gtacagtcgg cattatctca tattataaaa 1740
gccagtcatt aggcctatct gacaattcct gaatagagtt cataaacaat cctgcatgat 1800
aaccatcaca aacagaatga tgtacctgta aagatagcgg taaatatatt gaattacctt 1860
tattaatgaa ttttctgct gtaataatgg gtagaaggta attactatta ttattgatat 1920
ttaagttaaa ccagtaaat gaagtccatg gaataataga aagagaaaaa gcattttcag 1980
gtataggtgt tttgggaaac aatttccccg aaccattata tttctctaca tcagaaagg 2040
ataaatcata aaactctttg aagtcattct ttacaggagt ccaaatacca gagaatgttt 2100
tagatacacc atcaaaaatt gtataaagtg gctctaactt atcccaataa cctaactctc 2160
cgtcgctatt gtaaccagtt ctaaaagctg tatttgagtt tatcaccctt gtcactaaga 2220
aaataaatgc agggtaaaat ttatatcctt cttgttttat gtttcggtat aaaacactaa 2280

```

tatcaatttc	tgtggttata	ctaaaagtgc	tttgttggtt	caaataatga	ttaaatatct	2340
cttttctctt	ccaattgtct	aaatcaattt	tattaaagtt	catttgatat	gcctcctaaa	2400
tttttatcta	aagtgaattt	aggaggctta	cttgtctgct	ttcttcatta	gaatcaatcc	2460
ttttttaaaa	gtcaatatta	ctgtaacata	aatatatatt	ttaaaaatat	cccactttat	2520
ccaattttcg	tttgttgaac	taatgggtgc	tttagttgaa	gaataaagac	cacattaaaa	2580
aatgtggtct	tttgtgtttt	tttaaaggat	ttgagcgtag	cgaaaaatcc	ttttctttct	2640
tatcttgata	ataagggtaa	ctattgcatg	ataagctgtc	aaacatgaga	attcccgttt	2700
tcttctgcaa	gccaaaaaac	cttccgttac	aacgagaagg	attcttcact	ttctaaagtt	2760
cggcgagttt	catccctctg	tcccagtcct	tttttgatc	aaggcagact	gctgcaatgt	2820
ctatctatct	taataatagg	tgcagttcgc	aggcgatact	gcccaatgga	agtataccaa	2880
aatcaacggg	cttgtaacca	cacattagcc	caattcgata	tgcgcagaat	agattttttt	2940
aatgccttcg	ttcgtttcta	aaagcagaac	gccttcacat	tctataccta	acgccttacc	3000
gtaaaagggt	ccgtttaacg	ttctggctct	catattagt	ccaataccga	gcgcatagct	3060
ttcccataaa	agcttaatcg	gcgtaaatcc	gtgcgtcata	taatcccgt	accgtttctc	3120
aaagcatagt	aaaatatgct	ggatgacgcc	ggcccgatca	attttttccc	cagcagcttg	3180
gctgaggctt	gtcgcgatgt	ccttcaattc	atctggaaaa	tcattaggct	gctggttaaa	3240
cggctctccag	cttggctggt	ttggcggatg	agagaagatt	ttcagcctga	tacagattaa	3300
atcagaacgc	agaagcggtc	tgataaaaaca	gaatttgctt	ggcggcagta	gcgcggtggt	3360
cccactgac	cccatgccga	actcagaagt	gaaacgccgt	agcgcgatg	gtagtgtggg	3420
gtctcccat	gcgagagtag	ggaactgcc	ggcatcaaat	aaaacgaaag	gctcagtcga	3480
aagactgggc	ctttcgtttt	atctgttgtt	tgtcggtgaa	cgctctcctg	agtaggacaa	3540
atccgccggg	agcggatttg	aacgttgcca	agcaacggcc	cggagggtgg	cgggcaggac	3600
gcccgccata	aactgccagg	catcaaatta	agcagaaggc	catcctgacg	gatggccttt	3660
ttgcgtttct	acaaactcct	tttgtttatt	tttctaaata	cattcaaata	tgtatccgct	3720
catgagacaa	taaccctgat	aaatgcttca	ataatattga	aaaaggaaga	gtatgagtat	3780
tcaacatttc	cgtgtcgccc	ttattccctt	ttttgcggca	ttttgccttc	ctgtttttgc	3840
tcacccagaa	acgctggtga	aagtaaaaaga	tgctgaagat	cagttgggtg	cacgagtggg	3900
ttacatcgaa	ctggatctca	acagcggtaa	gatccttgag	agttttcgcc	ccgaagaacg	3960
ttttccaatg	atgagcaact	ttaaagttct	gctatgtggc	gcggtattat	cccgtgttga	4020
cgccgggcaa	gagcaactcg	gtcgcgccat	acactattct	cagaatgact	tggttgagta	4080
ctcaccagtc	acagaaaagc	atcttacgga	tggcatgaca	gtaagagaat	tatgcagtgc	4140
tgccataacc	atgagtgata	acactgcggc	caacttactt	ctgacaacga	tcggaggacc	4200
gaaggagcta	accgcttttt	tgcacaacat	gggggatcat	gtaactcgcc	ttgatcggtg	4260
ggaaccggag	ctgaatgaag	ccataccaaa	cgacgagcgt	gacaccacga	tgccctgtagc	4320
aatggcaaca	acgttgcgca	aactattaac	tggcgaacta	cttactctag	cttcccggca	4380
acaattaata	gactggatgg	aggcggataa	agttgcagga	ccacttctgc	gctcggccct	4440
tccggctggc	tggtttattg	ctgataaaatc	tggagccggt	gagcgtgggt	ctcgcggtat	4500
cattgcagca	ctggggccag	atggtaagcc	ctcccgtatc	gtagtattct	acacgacggg	4560
gagtcaggca	actatggatg	aacgaaatag	acagatcgct	gagatagggt	cctcactgat	4620
taagcattgg	taactgtcag	accaagttta	ctcatatata	ctttagattg	atttaaaact	4680
tcatttttaa	tttaaaagga	tctaggtgaa	gatccttttt	gataatctca	tgaccaaaat	4740
cccttaacgt	gagttttcgt	tccactgagc	gtcagaccpc	gtagaaaaga	tcaaaggatc	4800
ttcttgagat	cctttttttc	tgcgcgtaat	ctgctgcttg	caaacaaaaa	aaccaccgct	4860
accagcggtg	gtttgtttgc	cggatcaaga	gctaccaact	ctttttccga	aggtaactgg	4920
cttcagcaga	gcgcagatac	caaatactgt	ccttctagt	tagccgtagt	taggccacca	4980
cttcaagaac	tctgtagcac	cgcctacata	cctcgctctg	ctaactcctgt	taccagtggc	5040
tgctgccagt	ggcgataagt	cgtgtcttac	cgggttgac	tcaagacgat	agttaccgga	5100
taaggcgagc	cggtcgggct	gaacgggggg	ttcgtgcaca	cagcccagct	tggaagcaac	5160
gacctacacc	gaactgagat	acctacagcg	tgagctatga	gaaagcgcca	cgtttcccga	5220
agggagaaag	gcggagagg	atccggtta	cggcagggtc	ggaacaggag	agcgcacgag	5280
ggagcttcca	gggggaaacg	cctgggtatct	ttatagtctt	gtcgggtttc	gccacctctg	5340
acttgagcgt	cgatttttgt	gatgctcgtc	agggggcg	agcctatgga	aaaacgccag	5400
caacgcggcc	tttttacggg	tcctggcctt	ttgctggcct	tttgctcaca	tggtctttcc	5460
tgcggttatcc	cctgattctg	tggataaccg	tattaccgcc	tttgagttag	ctgataaccgc	5520
tcgcccagc	cgaacgaccg	agcgcagcga	gtcagttagc	gaggaagcgg	aagagcgctt	5580
gatgcggtat	tttctcctta	cgcattctgt	cggattttca	caccgcata	ggtgcactct	5640
cagtacaatc	tgtctgatg	ccgcatagtt	aagccagtat	acactccgct	atcgctacgt	5700
gactgggtca	tggttgccgc	ccgacacccg	ccaacacccg	ctgacgcgcc	ctgacgggct	5760
tgctgtctcc	cggcatccgc	ttacagacaa	gctgtgaccg	tctccgggag	ctgcatgtgt	5820
cagagggttt	caccgtcatc	accgaaacgc	gcgaggcagc	agatcaattc	gcgcgcgaag	5880

gcgaagcggc	atgcataatg	tgcctgtcaa	atggacgaag	cagggattct	gcaaacccta	5940
tgctactccg	tcaagccgtc	aattgtctga	ttcgttacca	attatgacaa	cttgacggct	6000
acatcattca	ctttttcttc	acaaccggca	cggaaactcg	tcgggctggc	cccgggtgat	6060
tttttaata	cccgcgagaa	atagagttag	tcgtcaaaac	caacattgcg	accgacgggtg	6120
gcgataggca	tccgggtggg	gctcaaaagc	agcttcgcct	ggctgatacg	ttggctcctcg	6180
cgccagctta	agacgcta	ccctaactgc	tggcggaaaa	gatgtgacag	acgcgacggc	6240
gacaagcaaa	catgctgtgc	gacgctggcg	atatcaaaat	tgctgtctgc	caggtgatcg	6300
ctgatgtact	gacaagcctc	gcgtaccgga	ttatccatcg	gtggatggag	cgactcggtta	6360
atcgcttcca	tgcgccgcag	taacaattgc	tcaagcagat	ttatcgccag	cagctccgaa	6420
tagcgccctt	ccccttgccc	ggcggttaatg	atttgcccaa	acaggtcgct	gaaatgcggc	6480
tggtgcgctt	catccgggcg	aaagaacccc	gtattggcaa	atattgacgg	ccagttaagc	6540
cattcatgcc	agtaggcg	cggacgaaag	taaacccact	ggtgatacca	ttcgcgagcc	6600
tccggatgac	gaccgtagt	atgaatctct	cctggcggga	acagcaaaat	atcaccgggt	6660
cggcaaacaa	attctcgtcc	ctgatttttc	accacccctt	gaccgcgaat	ggtgagattg	6720
agaatataac	ctttcattcc	cagcggtcgg	tcgataaaaa	aatcgagata	accgttggcc	6780
tcaatcggcg	ttaaacccgc	caccagatgg	gcattaaacg	agtatcccgg	cagcagggga	6840
tcattttg	cttcagccat	acttttcata	ctccccccat	tcagag		6886

<210> 21

<211> 7140

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: vector

<220>

<223> pAN620

<400> 21

tcggcgccg	cttcgtcgac	cgaaacagca	gttataaggc	atgaagctgt	ccgggtttttg	60
caaaagtggc	tgtgactgta	aaaagaaatc	gaaaaagacc	gttttgtgtg	aaaacgggtct	120
ttttgtttcc	ttttaaccaa	ctgccataac	tcgaggccta	cctagcttcc	aagaaagata	180
tcctaacagc	acaagagcgg	aaagatgttt	tgttctacat	ccagaacaac	ctctgctaaa	240
attcctgaaa	aattttgcaa	aaagtgtgtg	actttatcta	caaggtgtgg	tataataatc	300
ttaacaacag	caggacgctc	tagattagaa	aggaggattt	aaatatgaga	cagattactg	360
atattttcaca	gctgaaagaa	gccataaaac	aataccattc	agagggcaag	tcaatcggat	420
ttgttccgac	gatgggggtt	ctgcatgagg	ggcatttaac	cttagcagac	aaagcaagac	480
aagaaaacga	cgccgttatt	atgagtattt	ttgtgaatcc	tgcaacaattc	ggccctaattg	540
aagattttga	agcatatccg	cgcgatattg	agcgggatgc	agctcttgca	gaaaacgccg	600
gagtcgatat	tcttttttacg	ccagatgctc	atgatatgta	tcccgggtgaa	aagaatgtca	660
cgattcatgt	agaaagacgc	acagacgtgt	tatgcggg	ctcaagagaa	ggacattttg	720
acggggctcg	gatcgtactg	acgaagcttt	tcaatctagt	caagccgact	cgtgcctatt	780
tcggtttaaa	agatgcgcag	caggtagctg	ttgttgatgg	gttaatcagc	gacttcttca	840
tggatattga	attggttcct	gtcgatacgg	tcagagagga	agacggctta	gccaaaagct	900
ctcgcaatgt	atacttaaca	gctgaggaaa	gaaaagaagc	gcctaagctg	tatcgggccc	960
ttcaaacaag	tgcggaactt	gtccaagccg	gtgaaagaga	tcctgaagcg	gtgataaaaag	1020
ctgcaaaaga	tatcattgaa	acgactagcg	gaaccataga	ctatgtagag	ctttattcct	1080
atccggaact	cgagcctgtg	aatgaaattg	ctggaaagat	gattctcgct	gttgaggttg	1140
ctttttcaaa	agcgcgttta	atagataata	tcattattga	tattcgtaga	aaggaggtga	1200
atlaatatgt	atcgtacgat	gatgagcggc	aaacttcaca	gggcaactgt	tacggaagca	1260
aacctgaact	atgtgggaag	cattacaatt	gatgaagatc	tcattgatgc	tgtgggaatg	1320
cttcctaattg	aaaaagtaca	aattgtgaat	aataataatg	gagcacgtct	tgaaacgtat	1380
attattcctg	gtaaacgggg	aagcggcgctc	atatgcttaa	acgggtgcagc	cgcacgcctt	1440
gtgcaggaag	gagataagg	cattattatt	tcctacaaaa	tgatgtctga	tcaagaagcg	1500
gcaagccatg	agccgaaagt	ggctgttctg	aatgatcaaa	acaaaattga	acaaatgctg	1560
gggaacgaac	cagcccgtac	aattttgtaa	aggatcctgt	tttggcggtg	gagagaagat	1620
tttcagcctg	atacagatta	aatcagaacg	cagaagcggg	ctgataaaac	agaatttgcc	1680
tggcggcagt	agcgcgggtg	tcccacctga	ccccatgccg	aactcagaag	tgaacgccg	1740
tagcgccgat	ggtagtgtgg	ggtctcccca	tgcgagagta	gggaactgcc	aggcatcaaa	1800

taaaacgaaa	ggctcagtcg	aaagactggg	cctttcgttt	tatctgttgt	ttgtcgggtga	1860
acgtctcct	gagtaggaca	aatccgccgg	gagcggattt	gaacgttgcg	aagcaacggc	1920
ccggagggtg	gcgggcagga	cgcccgccat	aaactgccag	gcatcaaatt	aagcagaagg	1980
ccatcctgac	ggatggcctt	tttgcgtttc	tacaaactct	tggtaccgag	acgatcgtcc	2040
tctttgttgt	agcccatcac	ttttgctgaa	gagtaggagc	cgaaagtgac	ggcgtattca	2100
ttgagcggca	gctgagtcgc	accgacagaa	atcgcttctc	ttgatgtgcc	cggcgatccg	2160
actgtccagc	cgttcgggtc	gctgttgccg	tttgaggtaa	cagcgacaac	gccttctgac	2220
atggcccagt	caagcgctgt	gcttgctgcc	cagtcgggtt	tgtttaaaga	gtttccgaga	2280
gacaggttca	tcacatctgc	cccgtcctgc	actgcacgtt	ccacgccgcg	gatgacgttt	2340
tccgttgtgc	cgcttcggcc	aggccctaac	acacgataag	caagaagtgt	ggcatcaggc	2400
gctacgcctt	taatcgttcc	gtttgcagcc	acagttccgg	ctacgtgtgt	gccatggtca	2460
gttgccctgc	ccctcggatc	gccggttggg	gtttcttttg	gatcgtaatc	attgtccaca	2520
aaatcgtatc	ctttatattg	tccaaagttt	ttcttcagat	ctgggtgatt	gtattcaacc	2580
ccagtgtcaa	taatcgccac	cttgatgcct	tttctgtgtg	agcctaatac	ccatgcacgc	2640
tttgctccga	tataaggcgc	actgtcatcc	atttgcggag	atacggcgtc	ttcggagatt	2700
gtggggaatt	ctcatgtttg	acagcttatt	atgcaatagt	tacccttatt	atcaagataa	2760
gaaagaaaag	gatttttcgc	tacgctcaaa	tcctttaaaa	aaacacaaaa	gaccacattt	2820
tttaagtgtg	tctttattct	tcaactaaag	caccatttag	ttcaacaaac	gaaaattgga	2880
taaagtggga	tatttttaaa	atatatatatt	atgttacagt	aattattgact	tttaaaaaag	2940
gattgattct	aatgaagaaa	gcagacaagt	aagcctccta	aattcacttt	agataaaaaat	3000
ttaggaggca	tatcaaatga	actttaataa	aattgattta	gacaattgga	agagaaaaga	3060
gatattttaat	cattattttga	accaacaaac	gacttttagt	ataaccacag	aaattgatata	3120
tagtgtttta	taccgaaaca	taaaacaaga	aggatataaa	ttttaccctg	catttatttt	3180
cttagtgaca	aggggtgata	actcaaatac	agcttttaga	actgggttaca	atagcgacgg	3240
agagttaggt	tattgggata	agttagagcc	actttatata	atttttgatg	gtgtatctaa	3300
aacattctct	ggtattttgga	ctcctgtaaa	gaatgacttc	aaagagtttt	atgatttata	3360
cctttctgat	gtagagaaat	ataatgggtc	ggggaaattg	tttcccaaaa	cacctatacc	3420
tgaaaatgct	ttttctcttt	ctattattcc	atggacttca	tttactgggt	tttaacttaa	3480
tatcaataat	aatagtaatt	accttctacc	cattattaca	gcaggaaaat	tcattaataa	3540
aggtaattca	atatattttac	cgctatcttt	acaggtagat	cattctgttt	gtgatgggtta	3600
tcattgcagga	ttgtttatga	actctattca	ggaattgtca	gataggccta	atgactggct	3660
tttataatat	gagataatgc	cgactgtact	ttttacagtc	gggttttctaa	tgctactaac	3720
ctgccccgtt	agttgaagaa	cgaagcggcc	gcaattcttg	aagacgaaag	ggcctcgtga	3780
tacgcctatt	tttatagggt	aatgtcatga	taataatggg	ttcttagacg	tcagggtggca	3840
ccttttcggg	aaatgtgcgc	ggaaccccta	tttgtttatt	tttctaaata	cattcaaata	3900
tgatatccgt	catgagacaa	taaccctgat	aaatgcttca	ataatattga	aaaagggaaga	3960
gtatgagtat	tcaacatttc	cgtgtcggcc	ttattccctt	ttttgcggca	ttttgccttc	4020
ctgtttttgc	tcaccacagaa	acgtcgggtg	aagtaaaaga	tgctgaagat	cagttgggtg	4080
cacgagtggg	ttacatcgaa	ctggatctca	acagcggtaa	gatccttgag	agttttcgcc	4140
ccgaagaacg	ttttccaatg	atgagcactt	ttaaagttct	gctatgtggc	gcggtattat	4200
cccgtattga	cgccgggcaa	gagcaactcg	gtcgcgcgat	acactattct	cagaatgact	4260
tggttgagta	ctcaccagtc	acagaaaagc	atcttacgga	tggtcatgaca	gtaagagaat	4320
tatgcagtgc	tgccataacc	atgagtata	acactgcggc	caacttactt	ctgacaacga	4380
tcggaggacc	gaaggagcta	accgcttttt	tgcaacaacat	gggggatcat	gtaactcgcc	4440
ttgatcgttg	ggaaccggag	ctgaatgaag	ccataccaaa	cgacgagcgt	gacaccacga	4500
tgctgcagc	aatggcaaca	acgttcgcga	aactattaac	tggcgaacta	cttactctag	4560
cttcccggca	acaattaata	gactggatgg	aggcggataa	agttgcagga	ccacttctgc	4620
gctcggccct	tccggtcggc	tggtttattg	ctgataaaatc	tgagccgggt	gagcgtgggt	4680
ctcgcgggat	cattgcagca	ctggggccag	atggtaagcc	ctcccgtatc	gtagtattct	4740
acacgacggg	gagtcaggca	actatggatg	aacgaaatag	acagatcgct	gagatagggt	4800
cctcactgat	taagcattgg	taactgtcag	accaagttta	ctcatatata	cttttagattg	4860
atttaaaact	tcatttttaa	tttaaaagga	tctaggtgaa	gatccttttt	gataatctca	4920
tgaccaaaa	cccttaacgt	gagttttcgt	tccactgagc	gtcagacccc	gtagaaaaga	4980
tcaaaggatc	ttcttgagat	cctttttttc	tgccgcgtaat	ctgctgcttg	caaacaaaaa	5040
aaccaccgct	accagcggtg	gtttgtttgc	cggatcaaga	gctaccaact	ctttttccga	5100
aggtaactgg	cttcagcaga	gcgcagatac	caaatactgt	ccttctagt	tagccgtagt	5160
taggccacca	cttcaagaac	tctgtagcac	cgctacata	cctcgctctg	ctaactcctgt	5220
taccagtggc	tgctgccagt	ggcgataagt	cgtgtcttac	cgggttgagc	tcaagacgat	5280
agttaccgga	taaggcgcag	cggtcgggct	gaacgggggg	ttcgtgcaca	cagcccagct	5340
tgagcgcaac	gacctacacc	gaactgagat	acctacagcg	tgagctatga	gaaagcgcca	5400


```

cgcttcccga agggagaaaag gcggacaggt atccggtaag cggcagggtc ggaacaggag 5460
agcgacagag ggagcttcca gggggaacg cctggtatct ttatagtcct gtcgggtttc 5520
gccacctctg acttgagcgt cgatttttgt gatgctcgtc agggggcgcg agcctatgga 5580
aaaacgccag caacgcggcc tttttacggg tcttgccctt ttgctggcct tttgctcaca 5640
tgttctttcc tgcgttatcc cctgattctg tggataaccg tattaccgcc tttgagttag 5700
ctgataccgc tgcgcgcagc cgaacgaccg agcgacgcga gtcagtgage gaggaagcgg 5760
aagagcgctt gatgcgggat tttctcctta cgcattctgt cgggtatttca caccgcatat 5820
ggtgcactct cagtacaatc tgctctgatg ccgcatagtt aagccagtat acactccgct 5880
atcgctacgt gactgggtca tggctgcgcc ccgacacccg ccaacacccg ctgacgcgcc 5940
ctgacgggct tgtctgctcc cggcatccgc ttacagacaa gctgtgaccg tctccgggag 6000
ctgcatgtgt cagaggtttt caccgtcatc accgaaacgc gcgaggcagc tgcggtaaag 6060
ctcatcagcg tggctcgtgaa gcgattcaca gatgtctgcc tgttcatccg cgtccagctc 6120
gttgagtttc tccagaagcg ttaatgtctg gcttctgata aagcgggcca tgttaagggc 6180
ggttttttcc tgtttggtca cttgatgcct ccgtgtaagg ggaatttctt gttcatgggg 6240
gtaatgatac cgatgaaacg agagaggatg ctacagatac gggttactga tgatgaacat 6300
gcccggttac tggaaacgtt tgagggtaaa caactggcgg tatggatgcg gcgggaccag 6360
agaaaaatca ctacgggtca atgccagcgc ttcgtttaata cagatgtagg tgttccacag 6420
ggtagccagc agcatcctgc gatgcagatc cggaaacataa tgggtgcaggg cgctgacttc 6480
cgcgtttcca gactttacga aacacggaaa ccgaagacca ttcatgttgt tgcctcaggtc 6540
gcagacgttt tgcagcagca gtcgcttcac gttcgctcgc gtatcgggtg ttcatctctg 6600
taaccagtaa ggcaaccccc ccagcctagc cgggtcctca acgacaggag cagcatcatg 6660
cgcacccgtg gccaggaccc aacgctgcc gagatgcgcc gcgtgcggct gctggagatg 6720
gcggacgcga tggatatgtt ctgccaaggg ttggtttgcg cattcacagt tctccgaag 6780
aattgattgg ctccaattct tggagtgggt aatccgttag cgagggtgcc cgggcttcca 6840
ttcaggtcga ggtggcccgg ctccatgcac cgcgacgcaa cgcggggagg cagacaaggt 6900
atagggcggc gcctacaatc catgccaaac cgttccatgt gctcgcggag gcggcataaa 6960
tcgccgtgac gatcagcggg ccagtgatcg aagttagggt ggtaagagcc gcgagcgatc 7020
cttgaagctg tccctgatgg tcgtcatcta cctgcctgga cagcatggcc tgcaacgcgg 7080
gcatcccgat gccgccggaa gcgagaagaa tcataatggg gaaggccatc cagcctcgcg 7140

```

<210> 22

<211> 6725

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: vector

<220>

<223> pAN636

<400> 22

```

tcggcgccgc cttcgtcgac cgaaacagca gttataaggc atgaagctgt ccggtttttg 60
caaaagtggc tgtgactgta aaaagaaatc gaaaaagacc gttttgtgtg aaaacggtct 120
ttttgtttcc ttttaaccaa ctgccataac tcgaggccta cctagcttcc aagaaagata 180
tcctaacagc acaagagcgg aaagatgttt tgttctacat ccagaacaac ctctgctaaa 240
attcctgaaa aattttgcaa aaagtgtgtg actttatcta caaggtgtgg tataataatc 300
ttaacaacag caggacgctc tagaaggagg agaaaacatg aaaacaaaac tggattttct 360
aaaaatgaag gactctgaag aaccgattgt catgctgacc gcttatgatt atccggcagc 420
taaacttgct gaacaagcgg gagttgacat gatttttagtc ggtgattcac ttggaatggt 480
cgtcctcggc cttgattcaa ctgctgggtg gacagttgcg gacatgatcc atcatacaa 540
agccgttaaa aggggtgcgc cgaatacctt tattgtgaca gatatgccgt ttatgtctta 600
tcacctgtct aaggaagata cgctgaaaaa tgcagcggct atcgttcagg aaagcggagc 660
tgacgcactg aagcttgagg gcggagaagg cgtgtttgaa tccattcgcg cattgacgct 720
tgaggcgatt ccagtagtca gtcacttagg tttgacaccg cagtcatgct gcgtactggg 780
cggctataaa gtacagggca aagacgaaca aagcgccaaa aaattaatag aagacagtat 840
aaaatgcgaa gaagcaggag ctatgatgct tgtgctggaa tgtgtgccgg cagaactcac 900
agccaaaatt gccgagacgc taagcatacc ggtcattgga atcggggctg gtgtgaaagc 960
ggacggacaa gttctcgttt atcatgatat tatcgccac ggtgttgaga gaacaccta 1020
atttgtaaag caatatacgc gcattgatga aaccatcgaa acagcaatca gcggatatgt 1080

```

tcaggatgta	agacatcgtg	ctttccctga	acaaaagcat	tcctttcaaa	tgaaccagac	1140
agtgccttgac	ggcttgtacg	ggggaaaata	agggggggat	cctgttttgg	cggatgagag	1200
aagattttca	gcctgataca	gattaaatca	gaacgcagaa	gcggtctgat	aaaacagaat	1260
ttgcctggcg	gcagtagcgc	ggtgggtcca	cctgacccca	tgccgaactc	agaagtgaia	1320
cgccgtagcg	ccgatggtag	tgtggggctc	ccccatgcga	gagtagggaa	ctgccaggca	1380
tcaaataaaa	cgaaaggctc	agtcgaaaga	ctgggccttt	cgttttatct	gttgtttgtc	1440
ggtgaacgct	ctcctgagta	ggacaaatcc	gccgggagcg	gatttgaacg	ttgcgaagca	1500
acggcccggg	gggtggcggg	caggacgccc	gccataaact	gccaggcatc	aaattaagca	1560
gaaggccatc	ctgacggatg	gcctttttgc	gtttctacaa	actcttggtg	ccgagacgat	1620
cgtcctcttt	gttgtagccc	atcacttttg	ctgaagagta	ggagccgaaa	gtgacggcgt	1680
attcattgag	cggcagctga	gtcgcaccga	cagaaatcgc	ttctcttgat	gtgcccggcg	1740
atccgactgt	ccagccgttc	ggtccgctgt	tgccgtttga	ggtaacagcg	acaacgcctt	1800
ctgacatggc	ccagtcaagc	gctgtgcttg	tcgcccagtc	cgggttggtt	aaagagtttc	1860
cgagagacag	gttcatacaca	tctgcccgt	cctgcactgc	acgttccacg	cccgcgatga	1920
cgttttccgt	tgtgccgctt	ccgccaggcc	ctaaccacag	ataagcaaga	agtgtggcat	1980
caggcgctac	gccttttaac	gttccgtttg	cagccacagt	tcgggtacg	tgtgtgccat	2040
ggtcagttgc	ctcgcccctc	ggatcgccgg	ttgggtgttc	ttttggatcg	taatcattgt	2100
ccacaaaatc	gtatccttta	tattgtccaa	agtttttctt	cagatctggg	tgattgtatt	2160
caaccccgat	gtcaataatc	gccaccttga	tgccctttcc	tgtgtagcct	aaatcccatg	2220
catcgtttgc	tccgatataa	ggcgactgt	catccatttg	cggagatacg	gcgtcttcgg	2280
agattgtggg	gaattctcat	gtttgacagc	ttatcatgca	atagttaccc	ttattatcaa	2340
gataagaaaag	aaaaggattt	ttcgctacgc	tcaaatcctt	taaaaaaaca	caaaagacca	2400
catttttttaa	tgtggtcttt	attcttcaac	taaagcacc	attagttcaa	caaacgaaaa	2460
ttggataaaag	tgggataattt	ttaaaatata	tatttatgtt	acagtaatat	tgacttttaa	2520
aaaaggattg	atttctaata	agaaagcaga	caagtaagcc	tcctaaattc	actttagata	2580
aaaatttagg	aggcatatca	aatgaacttt	aataaaattg	atttagacaa	ttggaagaga	2640
aaagagatat	ttaatcatta	tttgaaccaa	caaacgactt	ttagtataac	cacagaaatt	2700
gatattagtg	ttttataccg	aaacataaaa	caagaaggat	ataaatttta	ccctgcattt	2760
attttcttag	tgacaagggt	gataaaactca	aatacagctt	ttagaactgg	ttacaatagc	2820
gacggagagt	taggttattg	ggataagtta	gagccacttt	atacaatttt	tgatgggtga	2880
tctaaaacat	tctctggtat	ttggactcct	gtaaagaatg	acttcaaaga	gttttatgat	2940
ttataccttt	ctgatgtaga	gaaatataat	gggttcggga	aattgtttcc	caaaacacct	3000
atacctgaaa	atgctttttc	tctttctatt	attccatgga	cttcatttac	tgggtttaac	3060
ttaaatatca	ataataatag	taattacctt	ctaccattta	ttacagcagg	aaaattcatt	3120
aataaaggta	attcaatata	tttaccgcta	tctttacagg	tacatcattc	tgtttgtgat	3180
ggttatcatg	caggattggt	tatgaactct	attcaggaat	tgtcagatag	gcctaataac	3240
tggtctttat	aatatgagat	aatgccgact	gtacttttta	cagtcgggtt	tctaatagtc	3300
ctaacctgcc	ccgttagttg	aagaacgaag	cggccgcaat	tcttgaagac	gaaagggcct	3360
cgtgatacgc	ctatttttat	aggttaatgt	catgataata	atgggtttct	agacgtcagg	3420
tggaactttt	cggggaaatg	tgcgcggaac	ccctattttg	ttatttttct	aaatacattc	3480
aaatatgtat	ccgctcatga	gacaataacc	ctgataaatg	cttcaataat	attgaaaaag	3540
gaagagtatg	agtattcaac	atttccgtgt	cgcccttatt	cccttttttg	cggcattttg	3600
ccttcctgtt	tttgctcacc	cagaaacgct	ggtgaaagta	aaagatgctg	aagatcagtt	3660
gggtgcacga	gtgggttaca	tcgaactgga	tctcaacagc	ggtaagatcc	ttgagagttt	3720
tcgccccgaa	gaacgttttc	caatgatgag	cactttttaa	gttctgctat	gtggcgcggt	3780
attatcccg	attgacgcgc	ggcaagagca	actcggctgc	cgcatacact	attctcagaa	3840
tgacttgggt	gagtactcac	cagtcacaga	aaagcatctt	acggatggca	tgacagtaag	3900
agaattatgc	agtgtcgcca	taaccatgag	tgataacact	gcggccaact	tacttctgac	3960
aacgatcgga	ggaccgaagg	agctaaccgc	ttttttgcac	aacatggggg	atcatgtaac	4020
tcgccttgat	cgttgggaac	cggagctgaa	tgaagccata	ccaaacgacg	agcgtgacac	4080
cacgatgccc	gcagcaatgg	caacaacggt	gcgcaacta	ttaactggcg	aactacttac	4140
tctagcttcc	cggcaacaat	taatagactg	gatggaggcg	gataaagttg	caggaccact	4200
tctgcgctcg	gcccttcogg	ctggctgggt	tattgctgat	aaatctggag	ccgggtgagcg	4260
tggtgtctgc	ggtatcattg	cagcactggg	gccagatggg	aagccctccc	gtatcgtagt	4320
tatctacacg	acggggagtc	aggcaactat	ggatgaacga	aatagacaga	tcgctgagat	4380
aggtgcctca	ctgattaagc	attggtaact	gtcagaccaa	gtttactcat	atatacttta	4440
gattgattta	aaacttcatt	tttaatttaa	aaggatctag	gtgaagatcc	tttttgataa	4500
tctcatgacc	aaaatccctt	aacgtgagtt	ttcggtccac	tgagcgtcag	accccgtaga	4560
aaagatcaaa	ggatcttctt	gagatccttt	ttttctgcgc	gtaatctgct	gcttgcaaac	4620
aaaaaaacca	ccgctaccag	cgggtggtttg	tttgccggat	caagagctac	caactctttt	4680

tccgaaggta	actggcttca	gcagagcgca	gataccaaat	actgtccttc	tagtgtagcc	4740
gtagttaggc	caccacttca	agaactctgt	agcaccgcct	acatacctcg	ctctgcta	4800
cctgttacc	gtggctgctg	ccagtggcga	taagtcgtgt	cttaccgggt	tggaactcaag	4860
acgatagtta	ccggataagg	cgcagcggtc	gggctgaacg	gggggttcgt	gcacacagcc	4920
cagcttgag	cgaacgacct	acaccgaact	gagataccta	cagcgtgagc	tatgagaaag	4980
cgccacgctt	cccgaaggga	gaaaggcgga	caggtatccg	gtaagcggca	gggtcggaac	5040
aggagagcgc	acgagggagc	ttccaggggg	aaacgcctgg	tatctttata	gtcctgtcgg	5100
gtttcgccac	ctctgacttg	agcgtcgatt	tttgtgatgc	tcgtcagggg	ggcggagcct	5160
atggaaaaac	gccagcaacg	cggccttttt	acggttcctg	gccttttgc	ggccttttgc	5220
tcacatgttc	tttctgctg	tatccccctga	ttctgtggat	aaccgtatta	ccgcctttga	5280
gtgagctgat	accgctcgcc	gcagccgaac	gaccgagcgc	agcgagtcag	tgagcgagga	5340
agcggaagag	cgctgatgc	ggtattttct	ccttacgcct	ctgtgcggta	tttcacaccg	5400
catatggtgc	actctcagta	caatctgctc	tgatgccgca	tagttaagcc	agtatacact	5460
ccgctatcgc	tacgtgactg	ggtcatggct	gcgccccgac	acccgccaac	acccgctgac	5520
gcgcccgtgac	gggcttgctc	gctcccggca	tcgcttaca	gacaagctgt	gaccgtctcc	5580
gggagctgca	tgtgtcagag	gttttcaccg	tcacaccga	aacgcgcgag	gcagctcg	5640
taaagctcat	cagcgtgggc	gtgaagcgat	tcacagatgt	ctgcctgttc	atccgcgtcc	5700
agctcggtga	gtttctccag	aagcgtta	gtctggcttc	tgataaagcg	ggccatgtta	5760
aggcggtttt	tttctgtttt	ggtcacttga	tgccctcgtg	taagggggaa	tttctgttca	5820
tgggggta	gataccgatg	aaacgagaga	ggatgctcac	gatacgggtt	actgatgatg	5880
aacatgccc	gttactggaa	cgttgtgagg	gtaaacaact	ggcggtatgg	atgcggcg	5940
accagagaaa	aatcactcag	ggtcaatgcc	agcgttctgt	taatacagat	gtaggtgttc	6000
cacagggtag	ccagcagcat	cctgcgatgc	agatccggaa	cataatggtg	cagggcgctg	6060
acttccgcgt	ttccagactt	tacgaaacac	ggaaaccgaa	gaccattcat	gttggtgtc	6120
aggtcgcaga	cgttttgcag	cagcagtcgc	ttcacgttcg	ctcgcgtatc	ggtgattcat	6180
tctgctaacc	agtaaggcaa	ccccgccagc	ctagccgggt	cctcaacgac	aggagcacga	6240
tcattgcgcac	ccgtggccag	gacccaacgc	tgcccagagat	gcgcccgcgtg	cggctgctgg	6300
agatggcgga	cgcatggat	atgttctgcc	aagggttgg	ttgcgcattc	acagttctcc	6360
gcaagaattg	attggctcca	attcttggag	tggtgaatcc	gttagcgagg	tgccgcggc	6420
ttccattcag	gtcgaggtgg	cccggctcca	tgaccgcga	cgcaacgcgg	ggaggcagac	6480
aaggatatagg	gcggcgccca	caatccatgc	caaccggttc	catgtgctcg	ccgaggcg	6540
ataaatcgcc	gtgacgatca	gcgggtccagt	gatcgaagtt	aggctggtaa	gagccgcgag	6600
cgatccttga	agctgtccct	gatggtcgctc	atctacctgc	ctggacagca	tggcctgcaa	6660
cgggggcatc	ccgatgccgc	cgggaagcgag	aagaatcata	atggggaagg	ccatccagcc	6720
tcgcg						6725

<210> 23

<211> 6806

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: vector

<220>

<223> pAN637

<400> 23

tcggcgccg	cttcgtcgac	cgaaacagca	gttataaggc	atgaagctgt	ccggttttt	60
caaaagtggc	tgtgactgta	aaaagaaatc	gaaaaagacc	gttttgtgtg	aaaacggctc	120
ttttgtttcc	ttttaaccaa	ctgccataac	tcgaggccta	cctagcttcc	aagaaagata	180
tcctaaccgc	acaagagcgg	aaagatgttt	tgttctacat	ccagaacaac	ctctgctaaa	240
attcctgaaa	aattttgcaa	aaagttgttg	actttatcta	caaggtgtgg	tataataatc	300
ttaacaacag	caggacgctc	tagacaattg	agatcttaag	aaaggaggtg	ttaattaatg	360
aagattggaa	tcattggcgg	aggetccgtt	ggtcttttat	gcgcctatta	tttgtcactt	420
tatcacgacg	tgactgttgt	gacgaggcgg	caagaacagg	ctgcggccat	tcagtctgaa	480
ggaatccggc	tttataaagg	cggggaggaa	ttcagggtcg	attgcagtgc	ggacacgagt	540
atcaattcgg	actttgacct	gcttgtcgctg	acagtgaagc	agcatcagct	tcaatctgtt	600
ttttcgtcgc	ttgaacgaat	cgggaagacg	aatatattat	ttttgcaaaa	cggcatgggg	660
catatccacg	acctaaaaga	ctggcacggtt	ggccattcca	tttatgttgg	aatcgttgag	720

cacggagctg	taagaaaatc	ggatacagct	gttgatcata	caggcctagg	tgcgataaaa	780
tggagcgctg	tgcacgatgc	tgaaccagac	cggctgaaca	tcttgtttca	gcataacccat	840
tcggattttc	cgattttatta	tgagacggat	tggtaccgtc	tgctgacggg	caagctgatt	900
gtaaatgcgt	gtattaatcc	tttaactgcg	ttattgcaag	tgaaaaatgg	agaactgctg	960
acaacgccag	cttatctggc	ttttatgaag	ctggtatttc	aggaggcatg	ccgcatttta	1020
aaacttgaaa	atgaagaaaa	ggcttgggag	cgggttcagg	ccgtttgtgg	gcaaacgaaa	1080
gagaatcggt	catcaatgct	ggttgacgtc	attggaggcc	ggcagacgga	agctgacgcc	1140
attatcggat	acttattgaa	ggaagcaagt	cttcaaggct	ttgatgccgt	ccacctagag	1200
tttttatatg	gcagcatcaa	agcattggag	cgaatatcca	acaaagtggg	ttactaagga	1260
tctgttttgg	gcggatgaga	gaagattttc	agcctgatac	agattaaatc	agaacgcaga	1320
agcggctctga	taaaacagaa	tttgccctggc	ggcagtagcg	cggtgggtccc	acctgacccc	1380
atgccgaact	cagaagtga	acgccgtagc	gccgatggta	gtgtgggggtc	tccccatgct	1440
agagtaggga	actgccaggc	atcaaataaa	acgaaaggct	cagtcgaaag	actgggcctt	1500
tcgtttttatc	tgttgtttgt	cggtgaacgc	tctcctgagt	aggacaaatc	cgccggggagc	1560
ggatttgaac	gttgcgaagc	aacggcccgg	agggtggcgg	gcaggacgcc	cgccataaac	1620
tgccaggcat	caaattaagc	agaaggccat	cctgacggat	ggcctttttg	cgtttctaca	1680
aactcttggg	accgagacga	tcgtcctctt	tgttgtagcc	catcactttt	gctgaagagt	1740
aggagccgaa	agtgcggcg	tattcattga	gcggcagctg	agtcgcaccg	acagaaatcg	1800
cttctcttga	tgtgcccggc	gatccgactg	tccagccgtt	cggtccgctg	ttgccgtttg	1860
aggtaacagc	gacaacgcct	tctgacatgg	cccagtcagg	cgctgtgctt	gtcgcccgagt	1920
ccgggttggt	taaagagttt	ccgagagaca	ggttcacac	atctgccccg	tctgacctg	1980
cacgttccac	gccccgcatg	acgttttccg	ttgtgccgct	tccgccaggc	cctaacacac	2040
gataagcaag	aagtgtggca	tcaggcgcta	cgcctttaat	cgttccgttt	gcagccacag	2100
ttccggctac	gtgtgtgcca	tggtcagttg	cctcgccccct	cggatcgccg	gttgggtgtt	2160
cttttgatc	gtaatcattg	tccacaaaat	cgtatccttt	atattgtcca	aagtttttct	2220
tcagatctgg	gtgattgtat	tcaaccccag	tgtcaataat	cgccaccttg	atgccttttc	2280
ctgtgtagcc	taaatcccat	gcacgttttg	ctccgatata	aggcgccactg	tcattccattt	2340
gcgagatac	ggcgtcttcg	gagattgtgg	ggaattctca	tgtttgacag	cttatcatgc	2400
aatagttacc	cttattatca	agataagaaa	gaaaaggatt	tttcgctacg	ctcaaactct	2460
ttaaaaaac	acaaaagacc	acatttttta	atgtggtcct	tattcttcaa	ctaaagcacc	2520
cattagttca	acaaacgaaa	attggataaa	gtgggatatt	tttaaaatat	atattttatgt	2580
tacagtaata	ttgactttta	aaaaaggatt	gattctaattg	aagaaagcag	acaagtaagc	2640
ctcctaaatt	cacttttagat	aaaaatttag	gaggcatatc	aaatgaactt	taataaaatt	2700
gatttagaca	attggaagag	aaaagagata	tttaatcatt	atltgaacca	acaaacgact	2760
tttagtataa	ccacagaaat	tgatattagt	gttttatacc	gaaacataaa	acaagaagga	2820
tataaatttt	accctgcatt	tattttctta	gtgacaaggg	tgataaactc	aaatacagct	2880
tttagaactg	gttacaatag	cgacggagag	ttaggttatt	gggataagtt	agagccactt	2940
tatacaattt	ttgatgggtg	atctaaaaca	ttctctggta	tttggactcc	tgtaaagaat	3000
gacttcaaag	agttttatga	tttatacctt	tctgatgtag	agaaatataa	tggttcgggg	3060
aaattgtttc	ccaaaacacc	tatacctgaa	aatgcttttt	ctctttctat	tattccatgg	3120
acttcattta	ctgggtttta	cttaaatatc	aataataata	gtaattacct	tctacccatt	3180
attacagcag	gaaaattcat	taataaagg	aattcaatat	atltaccgct	atctttacag	3240
gtacatcatt	ctgtttgtga	tggttatcat	gcaggattgt	ttatgaactc	tattcaggaa	3300
ttgtcagata	ggcctaata	ctggctttta	taatatgaga	taatgccgac	tgtacttttt	3360
acagtcggtt	ttctaattgc	actaacctgc	cccgttagtt	gaagaacgaa	gcggccgcaa	3420
ttcttgaaga	cgaaggggcc	tcgtgatacg	cctattttta	taggttaatg	tcatgataat	3480
aatggtttct	tagacgtcag	gtggcacttt	tcggggaaat	gtgcgcggaa	cccctattttg	3540
tttatttttc	taaatacatt	caaataatga	tccgctcatg	agacaataac	cctgataaat	3600
gcttcaataa	tattgaaaaa	ggaagagtat	gagtattcaa	catttccgtg	tcgcccttat	3660
tccctttttt	gcggcatttt	gccttcctgt	ttttgctcac	ccagaaacgc	tggtgaaagt	3720
aaaagatgct	gaagatcagt	tggtgtcacg	agtgggttac	atcgaaactg	atctcaacag	3780
cggtaaagtc	cttgagagtt	ttcgccccga	agaacgtttt	ccaatgatga	gcacttttta	3840
agttctgcta	tgtggcgcg	tattatcccg	tattgacgcc	gggcaagagc	aactcggtcg	3900
ccgcatacac	tattctcaga	atgacttggt	tgagtactca	ccagtcacag	aaaagcatct	3960
tacggatggc	atgacagtaa	gagaattatg	cagtgtgcc	ataaccatga	gtgataacac	4020
tgccggcaac	ttacttctga	caacgatcgg	aggaccgaag	gagctaaccg	cttttttgca	4080
caacatgggg	gatcatgtaa	ctcgccctga	tcgttgggaa	ccggagctga	atgaagccat	4140
accaaagcag	gagcgtgaca	ccacgatgcc	tgacgaatg	gcaacaacgt	tgcgcaaaact	4200
attaactggc	gaactactta	ctctagcttc	ccggcaacaa	ttaatagact	ggatggaggc	4260
ggataaagtt	gcaggaccac	ttctgcgctc	ggcccttccg	gctggctggt	ttattgctga	4320

```

taaactctgga gccggtgagc gtgggtctcg cggtatcatt gcagcactgg ggccagatgg 4380
taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg 4440
aaatagacag atcgtctgaga taggtgcctc actgattaag cattggtaac tgtcagacca 4500
agtttactca tatatacttt agattgattt aaaacttcat ttttaattta aaaggatcta 4560
ggtgaagatc ctttttgata atctcatgac caaaatccct taacgtgagt tttcgttcca 4620
ctgagcgtca gaccccgtag aaaagatcaa aggatcttct tgagatcctt tttttctgcy 4680
cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca gcggtggttt gtttgccgga 4740
tcaagagcta ccaactcttt ttccgaaggt aactggcttc agcagagcgc agataccaaa 4800
tactgtcctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc 4860
tacatacctc gctctgctaa tcctgttacc agtggctgct gccagtggcg ataagtctg 4920
tcttaccggg ttggactcaa gacgatagtt accggataag gcgcagcggc cgggctgaac 4980
gggggggttc tgacacacag ccagcttgga gcgaacgacc tacaccgaac tgagatacct 5040
acagcgtgag ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggatatcc 5100
ggtaagcggc agggctcgaa caggagagcg cagcagggag cttccagggg gaaacgcctg 5160
gtatctttat agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat ttttgatg 5220
ctcgtcaggg gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct 5280
ggccttttgc tggccttttg ctcacatgtt ctttctgcy ttatcccctg attctgtgga 5340
taaccgtatt accgcctttg agtgagctga taccgctcgc cgagccgaa cgaccgagcg 5400
cagcagtgca gtgagcgagg aagcggaaga gcgcctgatg cggtattttc tccttacgca 5460
tctgtgcggt atttcacacc gcatatggtg cactctcagt acaatctgct ctgatgccgc 5520
atagttaagc cagtatacac tccgctatcg ctacgtgact gggtcattggc tgcgccccga 5580
caccgcgcaa caccgctga cgcgcctga cgggcttgct tgctcccgcc atccgcttac 5640
agacaagctg tgaccgtctc cgggagctgc atgtgtcaga ggttttcacc gtcacaccg 5700
aaacgcgcga ggcagctgcy gtaaagctca tcagcgtggt cgtgaagcga ttcacagatg 5760
tctgcctgtt catccgcgtc cagctcgttg agtttctcca gaagcgttaa tgtctggctt 5820
ctgataaagc gggccatgtt aagggcgtt ttttctgctt tggtcacttg atgcctccgt 5880
gtaaggggga atttctgttc atgggggtaa tgataccgat gaaacgagag aggatgctca 5940
cgatacgggt tactgatgat gaacatgccc ggttactgga acgttgtag ggtaaacac 6000
tgccggtatg gatgcggcgg gaccagagaa aaatcactca gggccaatgc cagcgttctg 6060
ttaatacaga tgtaggtgtt ccacagggtg gccagcagca tcctgcatg cagatccgga 6120
acataatggt gcagggcgtt gacttccgcy tttccagact ttacgaaaca cggaaaccga 6180
agaccattca tgttgttgct caggtcgcag acgttttgca gcagcagtcg cttcacgttc 6240
gctcgcgtat cggtgattca ttctgctaac cagtaaggca accccgccag cctagccggg 6300
tcctcaacga caggagcacg atcatgcgca cccgtggcca ggaccacaacg ctgcccgaga 6360
tgccgccgct gcggctgctg gagatggcgg acgcgatgga tatgttctgc caagggttg 6420
tttgcgcatc cacagttctc cgcaagaatt gattggctcc aattcttgga gtgggtgaatc 6480
cgtagcgag gtgccgccgg cttccattca ggtcgagggt gcccggtcc atgcaaccgc 6540
acgcaacgcg gggaggcaga caaggtatag ggccgcgcct acaatccatg ccaaccgctt 6600
ccatgtgctc gccgaggcgg cataaatcgc cgtgacgatc agcgggtccag tgatcgaagt 6660
taggctggtg agagccgcga gcgatccttg aagctgtccc tgatggtcgt catctacctg 6720
cctggacagc atggcctgca acgcgggcat cccgatgccg ccggaagcga gaagaatcat 6780
aatggggaag gccatccagc ctcgcy 6806

```

<210> 24

<211> 3867

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: vector

<400> 24

```

aagctttctc aagaagcgaa caagaaaaaa gaagagcaga ttaaaccagct tcaagagttt 60
gtcgctagat tcagcgccaa tgcgtctaaa tctaagcagg ctacatcaag aaagaaactt 120
ctcgaaaaaa tcacgctgga tgatattaaa ccgtcttccc gccgctatcc ttatgttaac 180
ttcacgccgg aacgggaaat cggaaatgat gttcttcgcy tggaaaggctt aacaaaaaca 240
atcgatggcg taaaggtgct tgacaatgtc agctttatta tgaatcgaga agataaaatt 300
gctttcactg gccgaaatga acttgctgtt actacgctgt ttaaaatcat ttccggggaa 360
atggaagctg acagcggaac gtttaaatgg ggtgttacca catctcaagc gtattttcca 420
aaagacaaca gcgaatatatt cgaaggcagt gatctgaacc ttgtagactg gcttcgcca 480

```

tattctccgc	acgaccaaag	tgagagcttt	ttacgcggtt	tcttaggacg	catgctgttc	540
tctggagaag	aagtcacaaa	aaaagcaaat	gtactttccg	ggggagaaaa	ggtccgctgt	600
atgctgtcga	aagcgatgct	ttctggcgcc	aatatittaa	ttttggatga	gccgaccaac	660
catttagacc	tagagtccat	tacagcgctc	aataacggct	taatcagctt	taaaggcgct	720
atgctgttta	cttcccata	ccatcagttt	gtgcagacca	ttgccaacag	aattatagaa	780
attacaccta	acggcatcgt	cgataagcaa	atgagctatg	acgagttcct	tgaaaatgct	840
gatgtgcaga	aaaaattgac	tgaactatac	gccgaataaa	aaagcagaga	tttctctgct	900
ttttttgata	cctaaatgtg	aaaggagatc	acaacatgaa	atTTTTggtt	gtcggagcag	960
gtggagtagg	cgggtatatt	ggcggacggc	tttcggagaa	aggaaatgat	gtgacatttc	1020
tcgtgcgcca	aaaacgagct	gagcagctta	aaaaaacggg	gcttgtcatc	catagtgaag	1080
aagggaatgt	atcatttcag	cccgaactaa	tcagtgccgg	agaaacaggg	caatttgatg	1140
tcgttatcat	tgcttctaaa	gcatactcgc	ttggtcaagt	gatagaggat	gtcaaaccat	1200
ttatccatca	agaatctgtc	attatccctt	ttttaaatgg	gtaccgccac	tatgagcagc	1260
tatttgccgc	attttcaaaa	gaacagggtg	tgggcggcct	gtgtttttata	gaaagtgtct	1320
tagacaacaa	aggagaaatt	catcatacga	gcgcacgcga	tcgtttttgta	tttgaggaaat	1380
ggaacggcga	gcgtacggag	cggataagag	cgcttgaaga	ggcatttttca	ggtgtgaagg	1440
ctgaagtcac	cattagcggg	catatcgaga	agatcccttg	cagcaatagt	tacccttatt	1500
atcaagataa	gaaagaaaag	gattttttcgc	tacgtcctaa	tccttttaaaa	aaacacaaaa	1560
gaccacattt	tttaattgtg	tcctttattct	tcaactaaag	caccatttag	ttcaacaaac	1620
gaaaattgga	taaagtggga	tattttttaaa	atatataatt	atgttacagt	aatattgact	1680
tttaaaaaag	gattgattct	aatgaagaaa	gcagacaagt	aagcctccta	aattcacttt	1740
agataaaaaat	ttaggaggca	tatcaaatga	actttaataa	aattgattta	gacaattgga	1800
agagaaaaaga	gatattttaat	cattattttga	accaacaaac	gacttttagt	ataaccacag	1860
aaattgatata	tagtgtttta	taccgaaaca	taaaacaaga	aggatataaa	ttttaccctg	1920
cattttatttt	cttagtgaca	aggggtgataa	actcaaatac	agctttttaga	actggttaca	1980
atagcgacgg	agagttaggt	tattgggata	agttagagcc	actttataca	atttttgatg	2040
gtgtatctaa	aacattctct	ggtatttgga	ctcctgtaaa	gaatgacttc	aaagagtttt	2100
atgatttata	cctttctgat	gtagagaaat	ataatggttc	ggggaaattg	tttcccaaaa	2160
cacctatacc	tgaaaatgct	ttttctcttt	ctattattcc	atggacttca	tttactgggt	2220
tttaacttaaa	tatcaataat	aatagtaatt	accttctacc	cattattaca	gcaggaaaat	2280
tcattaataa	aggtaattca	atatattttac	cgctatcttt	acaggtagat	cattctgttt	2340
gtgatgggta	tcagtcagga	ttgtttatga	actctattca	ggaattgtca	gataggccta	2400
atgactggct	tttataatat	gagataatgc	cgactgtact	ttttacagtc	ggttttctaa	2460
tgctactaac	ctgccccgtt	agttgaagaa	ggtttttata	ttacagctcc	cgggagatct	2520
gggatcacta	gtccaaacga	cagaaggcga	ccacctgcat	ggatttttga	ttgaaaaagc	2580
aaaacgttta	tctctcgctg	caccagtatt	agaaaccggt	tatgcgaatc	tgcaaatgta	2640
tgaagcagaa	aaataaaaaa	aggaggcgga	aaagcctcct	tttatttact	taaaaagccc	2700
aatttccggt	tctgaagata	ggctctcttt	tccgctctgc	cgtaattccg	tcaatattca	2760
tatccttaga	accgatcata	aagtccacgt	gtgtaatgct	ttcatttagg	ccttctttga	2820
caagctcttc	acgagacatc	tgctttccgc	cttcaatatt	aaaggcatag	gcgcttccga	2880
tcgccaaatg	atttgacgcg	ttttcatcaa	acagcgtggt	atagaaaaga	atgtttgatt	2940
gtgatatagg	cgaatcgtaa	ggaacaagtg	ccacttcacc	taaaatagtg	gaaccttcac	3000
ctgtttccac	cagttctttt	aaaatatcct	cacctttttc	agctttaatg	tcgactatac	3060
ggccattttc	aaacgtcagg	gtgaaatttt	caataatatt	tccgccgtag	cttaatgggt	3120
ttgtgcttga	taccactccg	tcaaccccg	ctttttgcgg	cagcgtgaac	acttcttctg	3180
tcggcatatt	ggccataaac	tcagggccac	tttcattcac	gcttcccgca	cctgccccaa	3240
catgcttctc	aggcagctta	attgttagat	cagttccttc	tgcttgataa	tgtaaggcag	3300
cgtaatgtct	ctcggtcaaa	tggtcaactt	tttcatgaag	atTTTtggtca	tgattgatcc	3360
acgcctgaac	cggtgtgtct	tcattttacgc	gcgtcgcttt	aaaaattttct	tcccacagaa	3420
ggtggatcgc	ttcctcctct	gattttgccag	gaaacacctt	gtgagcccag	cctgctgatg	3480
ccgcacctac	gacagtccag	ctgactttgt	ctgattgaat	atattgtctg	tatgtatgta	3540
atgctttgcc	tgctgctttt	tggaatgccg	caatccggtt	ggaatctata	ccttttagca	3600
agtctgggtt	cgacgacaca	acagaaatga	aagcagctcc	atTTTtggtca	agctcttctc	3660
tgctttttgc	ttcccatcca	ggatattctt	caaagcttcc	aaacggcgca	agttcgtatt	3720
ttaatttggc	gacttcgtca	tcctgccaat	tcacgggtgac	gttttttgcg	cccttttcat	3780
atgcgtgttt	tacaattaaa	cggacaaaat	cccgaacgtc	tgttgaagca	tttacgacta	3840
catactggcc	tttttgagca	ttaacgc				3867

<210> 25
 <211> 8704
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: vector

<220>
 <223> pAN238

<400> 25

```

gcggccgctt  cgctcgaccga  aacagcagtt  ataaggcatg  aagctgtccg  gtttttgcaa  60
aagtggctgt  gactgtaaaa  agaaatcgaa  aaagaccgtt  ttgtgtgaaa  acggtctttt  120
tgtttccttt  taaccaactg  ccataactcg  aggcctacct  agcttccaag  aaagatatcc  180
taacagcaca  agagcggaag  gatgttttgt  tctacatcca  gaacaacctc  tgctaaaatt  240
cctgaaaaat  ttgcaaaaa  gttgttgact  ttatctacaa  ggtgtggtat  aataatctta  300
acaacagcag  gacgctctag  aggaggagac  taacatgaaa  tttttggttg  tcggagcagg  360
tggtataggc  gggatatatt  gcggacggct  ttcggagaaa  ggaaatgatg  tgacatttct  420
cgtgcgccaa  aaacgagctg  agcagcttaa  aaaaaccggg  cttgtcatcc  atagtgaaaa  480
agggatgta  tcatttcagc  ccgaactaat  cagtgcgga  gaaacagggc  aatttgatgt  540
cgttatcatt  gcttctaaag  catactcgct  tgggtcaagt  atagaggatg  tcaaaccatt  600
tatccatcaa  gaatctgtca  ttatcccttt  tttaaatggg  taccgccact  atgagcagct  660
atgtgcgca  ttttcaaaag  aacaggtgct  gggcgccctg  tgttttatag  aaagtgcctt  720
agacaacaaa  ggagaaattc  atcatacgag  cgcacgcgat  cgttttgtat  ttggagaatg  780
gaacggcgag  cgtacggagc  ggataagagc  gcttgaagag  gcattttcag  gtgtgaaggc  840
tgaaatcatc  attagcgggc  atatcgagaa  ggacatttgg  aaaaagtatc  tctttattgc  900
agcgcaagcg  gggatcacaa  cgttatttca  acgaccgctt  ggcccaatcc  tcgccacaga  960
agccggacgt  cacacggccc  aaactcttat  tggggaaatt  tgcactgttt  tacgaaaaga  1020
aggtgttccg  gctgatccgg  ctcttgagga  agagagcttt  cgtacgatga  ccagcatgtc  1080
ttaccatatg  aagtcctcca  tgcttcggga  tatggaaaac  ggccaaacga  cagaaggcga  1140
ccacctgcat  ggatttttga  ttgaaaaagc  aaaacgttta  tctctcgtct  caccagtatt  1200
agaaaccgtt  tatgcgaatc  tgcaaatgta  tgaagcagaa  aaataaaaaa  aggaggcgga  1260
aaagcctcct  tttattttact  taaaaagccc  aattttccgt  tctgaagata  ggctctcttt  1320
tcccgtctgc  cgggacactg  ttttggcgga  tgagagaaga  ttttcagcct  gatacagatt  1380
aaatcagaac  gcagaagcgg  tctgataaaa  cagaatttgc  ctggcgccag  tagcgcggtg  1440
gtcccacctg  accccatgcc  gaactcagaa  gtgaaacgcc  gtacgcccga  tggtagtggt  1500
gggtctcccc  atgcgagagt  agggaaactg  caggcatcaa  ataaaacgaa  aggctcagtc  1560
gaaagactgg  gcctttcgtt  ttatctgttg  tttgtcgggt  aacgctctcc  tgagtaggac  1620
aaatccgccc  ggagcggatt  tgaacgttgc  gaagcaacgg  cccggagggt  ggcgggcagg  1680
acgcccgcga  taaactgcca  ggcatcaaat  taagcagaag  gccatcctga  cggatggcct  1740
ttttgcgttt  ctacaaactc  ttggtaacca  gaaaaagcgg  caaaagcggc  tggttaaaaa  1800
gcgaaatcga  agaagctgtc  tgccgctaag  acggaatatc  aaaagcgttc  tgctgttggt  1860
tcatctttaa  aagtcacagc  cgatgaatcc  cagcaagatg  tcctaaaata  cttgaacacc  1920
cagaaagata  aaggaaatgc  agaccaaaat  cattcttatt  atgtggtgaa  cgggattgct  1980
gttcatgcct  caaaagaggt  tatggaaaaa  gtggtgcagt  tcccgaagt  ggaaaagggt  2040
cttcctaata  agaaacggca  gctttttaag  tcatcctccc  catttaatat  gaaaaaagca  2100
cagaaagcta  ttaaagcaac  tgacggtgtg  gaatggaatg  tagaccaaat  cgatgcccc  2160
aaagcttggg  cacttggaat  tgatggaact  ggcacggttg  ttgcgtccat  tgataccggg  2220
gtggaatgga  atcatccggc  attaaaagag  aaatatcgcg  gatataatcc  ggaaaatcct  2280
aatgagcctg  aaaatgaaat  gaactggtat  gatgccgtag  caggcgaggc  aagcccttat  2340
gatgatttgg  tcatggaac  ccacgtgaca  ggcacgatgg  tgggctctga  acctgatgga  2400
acaaatcaaa  ctggtgtagc  acctggcgca  aaatggattg  ctgttaaagc  gttctctgaa  2460
gatggcgcca  ctgagctga  cattttggaa  gctggtgaat  gggtttttag  accaaaggac  2520
gcggaaggaa  atccccacct  ggaaatggct  cctgatgttg  tcaataactc  atggggaggg  2580
ggctctggac  ttgatgaatg  gtacagagac  atgggtcaat  cctggcggtc  ggccgatatt  2640
ttccctgagt  tttcagcggg  gaatacggat  ctctttattc  ccggcgggcc  tggttctatc  2700
gcaaatccgg  caaactatcc  agaatcggtt  gcaactggag  cgactgagaa  ttccaattcc  2760
ccatggagag  aaaagaaaat  cgctaattgt  gattactttg  aacttctgca  tattcttgaa  2820
tttaaaaagg  ctgaaagagt  aaaagattgt  gctgaaatat  tagagtataa  acaaaatcgt  2880

```


gaaacaggcg	aaagaaagtt	gtatcgagtg	tggttttcta	aatccaggct	ttgtccaatg	2940
tgcaactgga	ggagagcaat	gaaacatggc	attcagtcac	aaaaggttgt	tgctgaagtt	3000
attaaacaaa	agccaacagt	tcgttggttg	tttctcacat	taacagttaa	aaatgtttat	3060
gatggcgaag	aattaaataa	gagtttgtca	gatatggctc	aaggatttcg	ccgaatgatg	3120
caatataaaa	aaattaataa	aaatcttggt	ggttttatgc	gtgcaacgga	agtgacaata	3180
aataataaag	ataattctta	taatcagcac	atgcatgtat	tggtatgtgt	ggaaccaact	3240
tattttaaga	atacagaaaa	ctacgtgaat	caaaaacaat	ggattcaatt	ttggaaaaag	3300
gcaatgaaat	tagactatga	tccaaatgta	aaagttcaaa	tgattcgacc	gaaaaataaa	3360
tataaatcgg	atatacaatc	ggcaattgac	gaaactgcaa	aatatcctgt	aaaggatacg	3420
gattttatga	ccgatgatga	agaaaagaat	ttgaaacggt	tgtctgattt	ggaggaaggt	3480
ttacaccgta	aaaggttaat	ctcctatggg	ggtttggtta	aagaaataca	taaaaaatta	3540
aaccttgatg	acacagaaga	aggcgatttg	attcatcacg	atgatgacga	aaaagccgat	3600
gaagatggat	tttctattat	tgcaatgtgg	aattgggaac	ggaaaaatta	ttttattaaa	3660
gagtagttca	acaaacgggc	catattgttg	tataagtgat	gaaatactga	attttaaact	3720
tagtttataat	gtggtaaaat	gttttaataca	agtttaggag	gaattaatta	tgaagtgtaa	3780
tgaatgtaac	agggttcaat	taaaagaggg	aagcgtatca	ttaaccctat	aaactacgtc	3840
tgccctcatt	attggagggt	gaaatgtgaa	tacatcctat	tcacaatcga	atttacgaca	3900
caaccaaaat	ttaatttggc	tttgcaatttt	atcttttttt	agcgtattaa	atgaaatggg	3960
tttgaacgtc	tcattacctg	atattgcaaa	tgattttaat	aaaccacctg	cgagtacaaa	4020
ctgggtgaac	acagccttta	tgtaaacctt	ttccattgga	acagctgtat	atggaaagct	4080
atctgatcaa	ttaggcatca	aaaggttact	cctatttgga	attataataa	attgtttcgg	4140
gtcggtaaat	gggtttgttg	gccattcttt	cttttccctta	cttattatgg	ctcgttttat	4200
tcaaggggct	gggtgcagctg	catttccagc	actcgtaatg	ggtgtagttg	cgcgctatat	4260
tccaaaggaa	aataggggta	aagcatttgg	tcttattgga	tcgatagtag	ccatgggaga	4320
aggagtcggg	ccagcgattg	gtggaatgat	agcccattat	attcattggg	cctatcttct	4380
actcattcct	atgataacaa	ttatcactgt	tccgtttctt	atgaaattat	taaagaaaga	4440
agtaaggata	aaaggtcatt	ttgatataca	aggaattata	ctaattgtctg	taggcattgt	4500
attttttatg	ttgtttacaa	catcatatag	catttctttt	cttatcggtt	gcgtgctgtc	4560
attcctgata	tttgtaaaac	atatcaggaa	agtaacagat	ccttttggtg	atcccggatt	4620
agggaaaaat	atacctttta	tgattggagt	tctttgtggg	ggaattatat	ttggaacagt	4680
agcagggttt	gtctctatgg	ttccttatat	gatgaaagat	gttcaccagc	taagtactgc	4740
cgaaatcgga	agtgtaatga	ttttccctgg	aacaatgagt	gtcattattt	tcggctacat	4800
tggtgggata	cttggttgata	gaagagggtc	tttatacgtg	ttaaaccatg	gagttacatt	4860
tctttctgtt	agctttttta	ctgcttccct	tcttttagaa	acaacatcat	ggttcatgac	4920
aattataatc	gtatttgttt	taggtgggct	ttcggtcacc	aaaacagtta	tatcaacaat	4980
tgtttcaagt	agcttgaaac	agcagggaag	tggtgctgga	atgagtttgc	ttaactttac	5040
cagcttttta	tcagagggaa	cagggtattg	aattgtaggg	ggtttattat	ccataccctt	5100
acttgatcaa	aggttgttac	ctatggaagt	tgatcagtc	acttatctgt	atagtaattt	5160
gttattactt	ttttcaggaa	tcattgtcat	tagttggctg	gttaccttga	atgtatataa	5220
acattctcaa	agggatttct	aaatcgttta	gggatcaact	ttgggagaga	gttcaaaatt	5280
gatccttttt	ttataacagt	tcgaagcggc	cgcaattctt	gaagacgaaa	gggcctcgtg	5340
atacgcctat	ttttataggg	taatgtcatg	ataataatgg	tttcttagac	gtcagggtggc	5400
acttttcggg	gaaatgtgcg	cggaacccct	atttgtttat	ttttctaaat	acattcaaat	5460
atgtatccgc	tcattgagaca	ataaccctga	taaatgcttc	aataatattg	aaaaaggaag	5520
agtatgagta	ttcaacattt	ccgtgtcgcc	cttattccct	tttttgcggc	attttgcctt	5580
cctgtttttg	ctcaccacga	aacgctggtg	aaagtaaaag	atgctgaaga	tcagttgggt	5640
gcacgagtg	gttacatcga	actggatctc	aacagcggta	agatccttga	gagttttcgc	5700
cccgaagaac	gttttccaat	gatgagcaat	tttaaagtct	tgctatgtgg	cgcggtatta	5760
tcccgtattg	acgcggggca	agagcaactc	ggtcgcccga	tacactattc	tcagaatgac	5820
ttggttgagt	actcaccagt	cacagaaaag	catcttacgg	atggcatgac	agtaagagaa	5880
ttatgcagtg	ctgccataac	catgagtgat	aacactgcgg	ccaacttact	tctgacaacg	5940
atcgaggagc	cgaaggagct	aaccgctttt	ttgcacaaca	tgggggatca	tgtaactcgc	6000
cttgatcggt	gggaaccgga	gctgaatgaa	gccataccaa	acgacgagcg	tgacaccacg	6060
atgcctgcag	caatggcaac	aacgttgccg	aaactattaa	ctggcggaact	acttactcta	6120
gcttcccggc	aacaattaat	agactggatg	gaggcggata	aagttgcagg	accacttctg	6180
cgctcggccc	ttccggctgg	ctgggtttat	gctgataaat	ctggagccgg	tgagcgtggg	6240
tctcgcggta	tcattgcagc	actggggcca	gatggtaagc	cctcccgtat	cgtagttatc	6300
tacacgacgg	ggagtcaggc	aactatggat	gaacgaaata	gacagatcgc	tgagataggt	6360
gcctcactga	ttaagcattg	gtaactgtca	gaccaagttt	actcatatat	acttttagatt	6420
gattttaaac	ttcattttta	attttaaagg	atctaggtga	agatcccttt	tgataatctc	6480

atgaccaaaa	tcccttaacg	tgagttttcg	ttccactgag	cgtcagacce	cgtagaaaag	6540
atcaaaggat	cttcttgaga	tccttttttt	ctgcgcgtaa	tctgctgctt	gcaaacaaaa	6600
aaaccaccgc	taccagcggg	ggtttggttg	ccggatcaag	agctaccaac	tctttttccg	6660
aaggtaactg	gcttcagcag	agcgcagata	ccaaatactg	tccttctagt	gtagccgtag	6720
ttaggccacc	acttcaagaa	ctctgtagca	ccgcctacat	acctcgctct	gctaatacctg	6780
ttaccagtgg	ctgctgccag	tgggcgataag	tcggtgtctta	ccgggttgga	ctcaagacga	6840
tagttaccgg	ataaggcgca	gcggtcgggc	tgaacggggg	gttcgtgcac	acagcccagc	6900
ttggagcgaa	cgacctacac	cgaactgaga	tacctacagc	gtgagctatg	agaaagcgcc	6960
acgcttcccc	aagggagaaa	ggcggacagg	tatccggtaa	gcggcagggt	cggaacagga	7020
gagcgcacga	gggagcttcc	agggggaaac	gcctggatc	tttatagtc	tgtcgggttt	7080
cgccacctct	gacttgagcg	tcgatttttg	tgatgctcgt	caggggggcg	gagcctatgg	7140
aaaaacgcca	gcaacgcggc	cttttttacgg	ttcctggcct	tttgctggcc	ttttgctcac	7200
atgttctttc	ctgcgttate	ccctgattct	gtggataacc	gtattaccgc	ctttgagtga	7260
gctgataacc	ctcgcgcgag	ccgaacgacc	gagcgcagcg	agtcagttag	cgaggaagcg	7320
gaagagcgcc	tgatgcggta	ttttctcctt	acgcatctgt	gcggtatttc	acaccgcata	7380
tggtgcactc	tcagtacaat	ctgctctgat	gccgcatagt	taagccagta	tacactccgc	7440
tatcgctacg	tgactgggtc	atggctgcgc	cccgcacccc	gccaacaccc	gctgacgcgc	7500
cctgacgggc	ttgtctgctc	ccggcatccg	cttacagaca	agctgtgacc	gtctccggga	7560
gctgcatgtg	tcagaggttt	tcaccgtcat	caccgaaacg	cgcgaggcag	ctgcggtaaa	7620
gctcatcagc	gtggctcgtga	agcgattcac	agatgtctgc	ctgttcatcc	gcgtccagct	7680
cgttgagttt	ctccagaagc	gttaatgtct	ggcttctgat	aaagcggggc	atgttaaggg	7740
cggttttttc	ctgtttggtc	acttgatgcc	tccgtgtaag	ggggaatttc	tgttcatggg	7800
ggtaatgata	ccgatgaaac	gagagaggat	gctcacgata	cgggttactg	atgatgaaca	7860
tgcccggtta	ctggaacggt	gtgagggtaa	acaactggcg	gtatggatgc	ggcgggacca	7920
gagaaaaatc	actcagggtc	aatgccagcg	cttcgttaat	acagatgtag	gtgttccaca	7980
gggtagccag	cagcatcctg	cgatgcagat	ccggaacata	atgggtgcagg	gcgctgactt	8040
ccgcgtttcc	agactttacg	aaacacggaa	accgaagacc	attcatgttg	ttgctcaggt	8100
cgcagacggt	ttgcagcagc	agtcgcttca	cgttcgctcg	cgtatcgggtg	attcattctg	8160
ctaaccagta	aggcaacccc	gccagcctag	ccgggtcctc	aacgacagga	gcacgatcat	8220
gcgcacccgt	ggccaggacc	caacgctgcc	cgagatgcgc	cgctgcggc	tgctggagat	8280
ggcggacgcg	atggatatgt	tctgccaagg	gttggtttgc	gcattcacag	ttctccgcaa	8340
gaattgattg	gctccaattc	ttggagtggg	gaatccgtta	gcgaggtgcc	gccggcttcc	8400
attcaggtcg	aggtggcccg	gctccatgca	ccgcgacgca	acgcggggag	gcagacaagg	8460
tatagggcgg	cgcctacaat	ccatgccaac	ccgttccatg	tgctcgccga	ggcggcataa	8520
atcgccgtga	cgatcagcgg	tccagtgatc	gaagttaggc	tggttaagagc	cgcgagcgat	8580
ccttgaagct	gtccctgatg	gtcgtcatct	acctgcctgg	acagcatggc	ctgcaacgcg	8640
ggcatcccga	tgccgccgga	agcgagaaga	atcataatgg	ggaaggccat	ccagcctcgc	8700
gtcg						8704